MOLECULAR STRUCTURAL INSIGHTS OF POLYGLUTAMINE-RICH AMYLOID-LIKE FIBRILS USING UV RESONANCE RAMAN SPECTROSCOPY

by

David Punihaole

B.S. in Molecular Biology, University of Pittsburgh, 2009

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

David Punihaole

It was defended on

June 24, 2016

and approved by

Dr. Sanford A. Asher, Distinguished Professor of Chemistry, Department of Chemistry

Dr. Linda Jen-Jacobson, Professor, Department of Biological Sciences

Dr. Ronald B. Wetzel, Professor, Department of Structural Biology

Dr. Sunil Saxena, Associate Professor, Department of Chemistry

Dr. Sean Garrett-Roe, Assistant Professor, Department of Chemistry

Dissertation Director: Dr. Sanford A. Asher, Distinguished Professor of Chemistry,

Department of Chemistry

MOLECULAR STRUCTURAL INSIGHTS OF POLYGLUTAMINE-RICH AMYLOID-LIKE FIBRILS USING UV RESONANCE RAMAN SPECTROSCOPY

David Punihaole, PhD

University of Pittsburgh, 2016

There is currently little that is known about the structure of polyglutamine (polyQ) fibrils, which are involved in at least ten neurodegenerative diseases, including Huntington's. Given the difficulty of studying these aggregates, new and incisive biophysical methods need to be developed in order to obtain high-resolution structural information of polyQ and other amyloid-like fibrils. Here, we present our recent advances in UV resonance Raman (UVRR) spectroscopy that enable the elucidation of molecular-level structural information of amyloid-like fibrils. We show, for example, how the primary amide UVRR bands report on the local hydrogen bonding and dielectric environment of glutamine side chains. We also discuss a newly discovered spectroscopic marker, the Amide III^P vibration, which sensitively reports on the OCCC dihedral angle of glutamine (Gln) and asparagine (Asn) side chains. These and other spectroscopic markers are used to gain insights into the peptide backbone and side chain conformations of polyQ peptides in solution-state and in fibrils. Finally, we demonstrate how the structural information obtained from UVRR can be utilized to guide Molecular Dynamics simulations in order to obtain experimentally validated structural models of polyQ fibrils.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

PREFACE

I would like to thank my advisor, Prof. Sanford Asher, for his mentorship and inspiration during my graduate school years. Most PhD students are "handed" a project in graduate school. Under Sandy's guidance, however, I was allowed to craft my own research program. I am extremely grateful for this rare opportunity, as it afforded me not only complete ownership of my work, but it also rigorously trained me to think critically as a scientist. I deeply admire and respect Sandy's prowess as a scientist and his infectious enthusiasm towards research. I could not have worked for a better PI.

Over the years, I have collaborated extensively with several people. I would like to thank Dr. Zhenmin Hong and Prof. Nataliya Myshakina for performing DFT calculations. I would also like to thank Riley Workman and Prof. Jeffry Madura for conducting Molecular Dynamics simulations. I am also grateful to Ryan Jakubek and Elizabeth Dahlburg for their assistance in conducting experiments and critically discussing the data.

My lab mates and colleagues in the Department of Chemistry are a source of inspiration, friendship, and support. I am especially grateful to Jonathan Wert, Yangguang Ou, and Chino Cabalteja. I would also like to thank Sharon Mansfield, our group's administrative assistant, for all her kindness, help, and support.

I would like to acknowledge my thesis committee, particularly Prof. Linda Jen-Jacobson. As an undergraduate student, I took a course taught by Linda, which inspired me to switch career paths and to pursue science. Her course instilled in me a deep appreciation for applying biophysical methods to gain fundamental insights into biological phenomena. It was probably the single most important and influential course that I took at the University of Pittsburgh.

Finally, I would like to express gratitude to my close friends and family, particularly

my mom and grandma. Their hard work, support, and sacrifices over the years enabled me to become, not only the first person in my family to earn a bachelor's degree, but also a doctorate.

1.0 POLYGLUTAMINE-RICH PROTEINS AND AGGREGATION

1.1 CLINICOPATHOLOGY OF CAG REPEAT DISEASES

The abnormal expansion of DNA trinucleotide repeats that encode homopolymeric tracts of amino acids in proteins has been implicated in at least 16 diseases [1]. One of the most prevalent and devastating classes of trinucleotide diseases involve genomic expansions of CAG repeats that encode for polyQ tracts in proteins. At present, expansions in CAG repeats have been linked to at least 10 diseases [1–4] (Table 1.1), including several sinocerebellar ataxis and spinobulbar muscular atrophy. However, the most notable and prevalent CAG repeat disorder is Huntington's disease.

The predominant clinical symptoms of CAG repeat diseases are neurodegenerative in nature (Table 1.1). The characteristic symptoms of Huntington's disease, for example, are involuntary, jerky movements (chorea), cognitive impairments, mood swings, and behavioral changes that progressively worsen over time. In the case of spinobulbar muscular atrophy, additional non-neurological symptoms also manifest, including gynecomastia and sterility.

A common feature of these diseases is that the severity of the clinicopathological symptoms correlates with the length of the polyQ repeat expansion. Generally speaking, the larger the repeat length, the greater the disease morbidity and mortality rate. There is strong epidemiological evidence that indicates successive generations of families afflicted by CAG repeat diseases experience an earlier age of onset of disease symptoms [1, 5]. In all of these diseases, symptoms manifest when affected proteins possess polyQ tracts that exceed a threshold repeat length. For example, in Huntington's disease [4, 6], repeats between 17–30 Gln residues in the huntingtin are generally considered benign, whereas repeat lengths that exceed 36 Gln residues typically result in disease symptoms.

Table 1.1: Summary of Trinucleotide CAG Repeat Diseases[a]

Disease[b]	Gene Product	Normal Repeat Length	Pathological Repeat Length	Clinical Features
HD	Huntingtin	6–34	36–121	Chorea, dystonia, cognitive impairment, depression, psychiatric problems
SCA1	Ataxin1	6–44	39–82	Ataxia, cognitive impairments, slurred speech
SCA2	Ataxin2	15–24	32–200	Ataxia, decreased reflexes, infant variant with retinopathy
SCA3	Ataxin3	13–36	61–84	Ataxia, parkinsonism
SCA6	$CACNA1_A$	10–33	36–121	Ataxia, dysarthria, tremors
SCA7	Ataxin7	4–35	37–306	Ataxia, blindness, infant variant with cardiac failure
SCA17	TBP	25–42	47–63	Ataxia, cognitive decline, seizures, psychiatric problems
SBMA	Androgen receptor	38–62	36–121	Gynecomastia, decreased fertility, motor weakness
DRPLA	Atrophin	7–34	49–88	Ataxia, seizures, dementia, choreoathetosis

[a]Table adapted from references [1, 3]. [b]HD: Huntington's disease; SCA: Spinocerebellar ataxia; SBMA: Spinobulbar muscular atrophy (Kennedy disease); DRPLA: Dentatorubral-pallidoluysian atrophy.

1.2 THE ROLE OF POLYQ-RICH AGGREGATES IN PATHOLOGY

The pathological hallmarks of all CAG repeat diseases is the accumulation of insoluble polyQ-rich protein aggregates in cells [7]. An important aspect in the pathophysiology of CAG repeat diseases is understanding the role that polyQ-rich aggregates play in neurotoxicity. Although these aggregates are pathological hallmarks, the question remains: Are they the toxic agents or merely by-products of the diseases? Numerous studies have addressed this question in recent years; however, there is little consensus, and the emerging picture is complex. Some studies link polyQ-rich aggregates with cellular toxicity and neurodegeneration, while others suggest that they potentially play a neuroprotective role.

1.2.1 Evidence Supporting the Toxic Role of Aggregates

Evidence that supports the toxic aggregate hypothesis stems from several observations. Studies on Huntington's disease, for example, show that transcription factors [8–12], molecular chaperones [13], and proteasomal proteins [14] colocalize with nuclear and cytoplasmic inclusions. These results suggest that polyQ aggregates may disrupt normal homeostasis by sequestering transcription factors and other proteins, thereby depleting their cellular concentrations and disrupting their normal functions [15, 16].

More direct evidence supporting aggregate toxicity derives from cell-based assays. Exogenously introduced polyQ aggregates are highly cytotoxic when localized to the nucleus [17]. In addition, expression of atrophin-1 with long tracts of polyQ repeats (129 Gln residues) in mice results in significantly more neuronal inclusions and greater brain atrophy than protein constructs with smaller repeat lengths [18]. These results mirror some postmortem studies [19, 20] of presymptomatic Huntington's patients, where the formation of polyQ-rich aggregates correlate to morphological changes in brain tissue.

The cytotoxic effects of nuclear inclusions can be mitigated by rescuing cells with inhibitors and molecular chaperones that can prevent aggregation. For example, Cummings *et al.* [21] showed that the overexpression of molecular chaperone proteins significantly decreases the accumulation of ataxin-1 aggregates. The most compelling evidence, however,

stems from Thakur *et al.* [22], who showed that treating cells with small proline-containing peptides that inhibit aggregation results in significantly reduced cell death.

1.2.2 Evidence Supporting the Non-Toxic Role of Aggregates

Despite mounting evidence, there are several studies that have strongly questioned whether polyQ-rich aggregates are the toxic agents. An early study that questioned the toxic aggregate hypothesis was published by Greenberg and coworkers [23]. Although their results showed that preventing huntingtin aggregation in the nucleus stopped apoptosis, they did not find a strong correlation between the formation of nuclear inclusions and cell death. Greenberg and coworkers noted, for example, that the formation of huntingtin aggregates accompanied neuronal cell death in the striatum, but not the hippocampus. Their findings have been corroborated by more recent studies [24, 25] that also fail to find a robust link between polyQ aggregates and cell death or neurodegenerative disease symptoms.

Other studies suggest that polyQ-rich aggregates may actually serve a neuroprotective role! Evidence supporting this hypothesis derives primarily from mice studies. For example, Cummings *et al.* [26] showed that reducing ataxin-1 aggregation acutally correlated to an increase in the disease pathology spinocerebellar ataxia-1. Their findings are supported by several more recent studies [27–29], which show that the most apoptotically vulnerable neuronal cells do not show the formation of aggregate inclusions until the most advanced stages of disease. In light of these studies, it has been proposed [1] that monomeric (and not aggregated) polyQ-rich proteins are the cytotoxic agents (*vide infra*).

1.2.3 Reconciling Contradictory Results on Aggregate Toxicity

It is difficult to reconcile the seemingly contradictory observations regarding the cellular toxicity of polyQ aggregates. There are many potential reasons for the conflicting reports, including significant differences in experimental designs and controls. It is conceivable, for instance, that the different cell lines and animal models used by various research groups exhibit differential susceptibilities to polyQ aggregate toxicity. In addition, the structural nature and location of polyQ aggregates may also be important. For example, aggregates

located in the nucleus are significantly more cytotoxic than those located in the cytoplasm [17]. Similarly, some studies even suggest that various aggregate polymorphs exhibit different cytotoxic effects on cells [30]. Finally, it may also be the case that the pathophysiologies of different CAG repeat diseases vary so that in some disorders aggregates play a more cytotoxic role. However, it is difficult to imagine that the formation of large neuronal polyQ-rich aggregates do not at least play some significant role in neuronal cell death.

1.3 UNDERSTANDING THE AGGREGATION OF POLYQ-RICH PEPTIDES AND PROTEINS

Understanding the underlying aggregation mechanisms of polyQ peptides and proteins is important in developing fundamental insights into the etiologies of CAG repeat diseases, as well as formulating potential therapies. To this end, biophysical studies have focused largely on understanding the structures of polyQ-rich peptides in both solution-state and in aggregates. Given the only apparent commonality of CAG repeat diseases are the polyQ repeats, most studies have focused on investigating synthetic homopolymeric model peptide systems, with sequences such as $D_2Q_nK_2$ or $K_2Q_nK_2$ (where n is the number of Gln repeats).

1.3.1 PolyQ Peptides are Disordered in Solution-state

Most biophysical studies indicate that polyQ peptides are structurally disordered in solution-state [31–35]. One of the most surprising findings is that there is no obvious difference in the solution-state secondary structures between "pathologic" and "non-pathologic" polyQ repeat lengths [34]. These findings are also supported by computational studies [36–39], which show that polyQ peptides, regardless of repeat length, are largely disordered, with only transient elements of regular secondary structures such as α-helices, β-sheets, and turns.

Although polyQ peptides are structurally disordered, they do not behave as true random coil polymers [4]. Several experimental and computational studies [36, 37, 40–44] indicate

that the end-to-end distances of polyQ peptides deviate significantly from random coils, since they adopt relatively collapsed, globule-like structures. In addition, a detailed analysis of circular dichroism (CD) and nuclear magnetic resonance (NMR) spectra [33] shows that that polyQ peptides show a high propensity towards adopting residual polyproline II-like (PPII-like) secondary structures, although not necessarily in long, continuous tracts.

These studies indicate that water is a poor solvent for polyQ peptides that contain long repeat lengths [43], despite the apparent hydrophilic nature of the Gln side chains. The collapsed, globule-like structures adopted by polyQ peptides are the result of the extensive hydrogen bonding interactions that can occur between the secondary amides of the peptide backbone and the primary amides of the Gln side chains. These inter-amide interactions are presumably stronger than amide-water hydrogen bonding interactions.

1.3.2 The Toxic Monomer Hypothesis

Although polyQ monomers are predominately disordered in solution-state, there are some studies that suggest these peptides adopt low concentrations of non-disordered conformations. These non-disordered structures are believed to be the putative cytotoxic agents that contribute to neurodegeneration [35]. According to this hypothesis, these non-disordered states increase significantly in concentration for pathological polyQ repeat lengths (*i.e.* 35–45 Gln residues in the case of Huntington's disease). The appeal of this so-called "toxic monomer hypothesis" is that it provides an explanation for the pathological gain-of-function that arises in mutated proteins with expanded polyQ repeat domains, while also accounting for observations that polyQ aggregates do not robustly correlate with neuronal cell death.

The supporting data for this hypothesis derives primarily from antibody binding studies. Trottier *et al.* [45], for example, observed with Western blot analysis that monoclonal anti-polyQ antibodies have a higher binding affinity for proteins with longer polyQ repeat domains than shorter, non-pathologically relevant repeat lengths. These data led to the speculation that the antibodies preferentially recognize stretches of polyQ repeats that adopt specific secondary structures. Similar results have been observed more recently with a different monoclonal antibody, called 3B5H10 [46]. X-ray crystallographic and small-angle X-ray

scattering data [47] suggests that 3B5H10 antibodies recognize β-hairpins that are formed in polyQ peptides with pathologically long repeat lengths.

Although it is difficult to unequivocally rule out the existence of low concentrations of polyQ monomers that adopt stable, ordered structures, there are alternative explanations that account for the results of these antibody binding studies. One explanation, for example, is that antibodies can trap kinetically accessible polyQ conformations that are not normally populated. Another explanation is the so-called "linear lattice" effect [48, 49], where the preferential binding is due to there being a significant increase in the number of epitopes in polyQ peptides of pathological repeat lengths compared to non-pathological repeat lengths.

1.3.3 Proposed Aggregate Structures

PolyQ aggregates prepared from synthetic peptides *in vitro* bear many of the hallmark features of amyloid-like fibrils, including displaying filamentous morphologies, exhibiting β-sheet-rich structures, and binding Thioflavin-T [32, 35]. Despite this, polyQ aggregates do not exhibit all the characteristics of amyloids as exemplified, for example, by Aβ fibrils. Some amyloid features that polyQ aggregates do not display include exhibiting birefringence upon Congo Red binding and a classical cross-β X-ray fiber diffraction pattern [4, 32, 35, 50]. There are currently no atomic-resolution structures of polyQ aymloid-like fibril aggregates. However, numerous structures (Figure 1.1) have been proposed for polyQ fibrils on the basis of a variety of classical biophysical methods.

The first detailed structural model of polyQ fibrils was proposed by Perutz and coworkers [52], who utilized X-ray fiber diffraction. Their results showed that $D_2Q_{15}K_2$ aggregates produce diffraction patterns with prominent reflections located at 4.2 Å, 4.8 Å, and 8.4 Å. They attributed the diffraction pattern to a cross-β structure and assigned the 4.8 Å reflection to the inter-strand distance of the fibril β-sheets. However, the origin of the 8.4 Å reflection was mysterious since typical amyloids show inter-β-sheet distances between 10–12 Å. As a result, Perutz *et al.* proposed that polyQ fibril aggregates were composed of "polar-zipper" structures (Figure 1.1a), where the neighboring β-sheets are stacked $\sim$17 Å apart and are hydrogen bonded to each other *via* side chain amides.

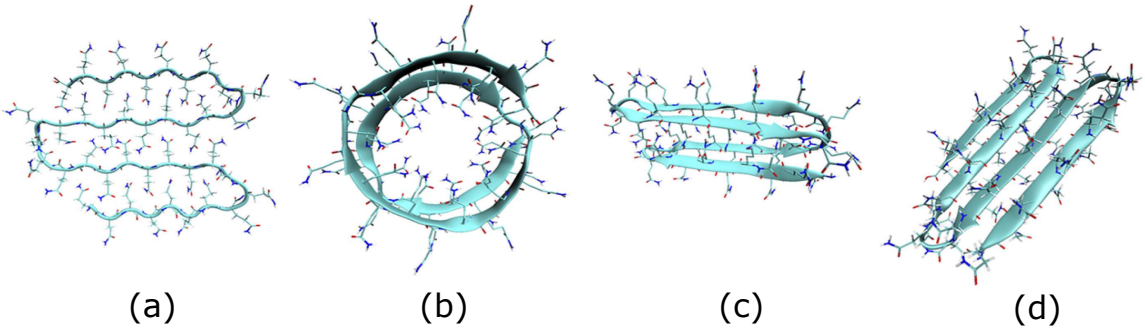


Figure 1.1: Proposed structures for polyQ fibrils: (a) polar zipper structure; (b) β-helix; (c) β-arc; and (d) canonical β-sheet. Adapted with permission from [51]. Copyright © (2012), American Chemical Society.

The lack of the characteristic 10–12 Å reflection in polyQ fibril diffraction patterns compelled Perutz and coworkers to reanalyze their data in a subsequent study [53]. In their revised model, they proposed that polyQ fibrils consist of water-filled nanotubes that are formed by cylindrical β-helices (Figure 1.1b). However, other studies provide alternative interpretations of the polyQ diffraction data collected by Perutz *et al.* [52, 53]. For example, Sikorski and Atkins [54] analyzed the original X-ray data of Perutz [52] and argued that the β-helical model was incorrect. Instead, they proposed that polyQ fibrils are composed of a cross-β structure, but, because of extensive side chain inter-amide hydrogen bonding interactions, the β-sheets are stacked closer together (∼8–9 Å) than typical amyloids (∼10–12 Å).

The Sikorski and Atkins interpretation of the Perutz data have been further substantiated by Sharma *et al.*'s [55] work, who also propose that polyQ fibrils are composed of cross-β structures. The Sharma *et al.* model posits that polyQ fibril aggregates are composed of antiparallel β-sheets that contain reverse hairpin turns. Their model, however, differs significantly from the Sikorski and Atkins [55] structure. For example, despite obtaining similar diffraction patterns as those analyzed by Sikorski and Atkins, Sharma *et al.* assign different unit cells and thus their proposed fibril structure contains significant structural differences.

More recent studies have investigated polyQ fibril structure using solid state NMR. Schneider *et al.* [56] studied a series of peptides with polyQ repeats ranging from 15–54

residues. They proposed that the basic structural motif of $D_2Q_{15}K_2$ fibrils is an extended β-strand structure, which assembles to form antiparallel β-sheets. Based on their interpretation of solid state NMR spectra of $GK_2Q_{38}K_2$ and $GK_2Q_{54}K_2$ aggregates, Schneider *et al.* propose that fibrils prepared from larger polyQ peptides adopt β-arc structures (Figure 1.1c), similar to those observed in Aβ [57].

The β-arc model has been challenged by other solid state NMR studies. Van der Wel and coworkers have measured solid state NMR spectra of polyQ fibrils prepared from both model and more complex peptides that contain flanking sequences found in the huntingtin protein [58–61]. Their fibril spectra are very similar to those reported by Schneider *et al.* [56]; however, van der Wel and coworkers believe the data are inconsistent with a β-arc structure. Instead, they hypothesize that polyQ fibrils are composed of canonical antiparallel β-sheet structures that contain reverse-hairpin turns [61] (Figure 1.1d).

There have been a number of computational studies that have examined the structure of polyQ fibrils. Early simulation studies suggested that β-helical nanotubes are potentially stable fibril structures [62–65]. However, these studies did not rigorously validate their simulation results against experiments. More recent and rigorous computational studies [51, 66] that employ multiple force fields clearly show that β-nanotubes are highly unstable. Instead, these studies indicate that the most stable fibril architectures are β-arc and β-sheet structures.

1.3.4 Proposed Aggregation Mechanisms

Mechanistic studies on polyQ fibril aggregation have been pioneered by Wetzel and coworkers [59, 67–71]. Using sedimentation assays [72], the Wetzel group [35] proposes that simple polyQ peptide fibril aggregation proceeds *via* a classical nucleated growth-polymerization model. Interestingly, their systematic studies examining the aggregation concentration dependence of $K_2Q_nK_2$ peptides indicate that polyQ fibrils are homogeneously nucleated and that the critical nucleus size is a monomer for peptides containing repeats of more than 25 Gln residues [32, 59].

Traditionally, a monomeric nucleus is considered unusual in polymer theory [32]. To

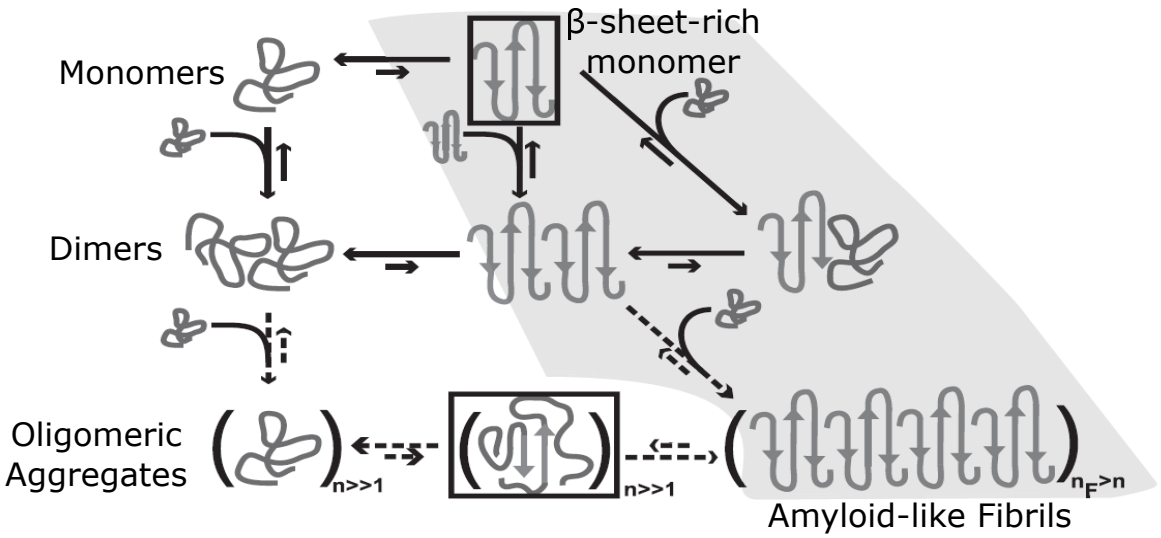


Figure 1.2: Proposed aggregation mechanisms for polyQ fibril formation. The region shaded in gray represents the aggregation pathway proposed by Wetzel and coworkers. In this pathway, nucleation occurs through the energetically unfavorable structural conversion of a structurally disordered monomeric polyQ peptide to a β-sheet-rich structure that can initiate fibril formation. The alternative pathway is proposed by Pappu and coworkers. In this mechanism, polyQ peptides oligomerize into globule-like aggregates. These aggregates contain interfaces that promote the formation of β-sheet structures that eventually lead to fibrils. This figure was reprinted and adapted from [38], copyright © (2009), with permission from Elsevier.

rationalize their observations, Wetzel and coworkers [32] proposed that fibril nucleation is initiated by the energetically unfavorable conversion of a structurally disordered polyQ peptide into a β-sheet-rich peptide through the formation of a hairpin turn (Figure 1.2). Wetzel and coworkers have published a number of elegant mutagenesis studies [59, 69] that provide evidence supporting this model. Their studies show, for example, that polyQ peptides containing β-hairpin enhancing motifs (*e.g.* tryptophan zippers, Pro-Gly insertions, Cysteine disulfide bonds, salt bridge forming residues) result in significantly increased aggregation kinetics compared to unmutated peptide sequences.

The classical nucleated growth-polymerization model has been challenged in recent years

by a series of computational studies conducted primarily by Pappu and coworkers [36–38, 41, 44, 62]. Their simulations corroborate [36, 41] experimental evidence that monomeric polyQ peptides are structurally disordered in solution-state. However, in contrast to the Wetzel model [73], Pappu and coworkers predict that β-sheet formation of monomeric polyQ peptides becomes more energetically unfavorable with increasing Gln repeat lengths. Their computational data [37, 38] indicates that the spontaneous formation of dimers and higher order oligomers increases dramatically as a function of polyQ repeat length due to the formation of non-specific hydrogen bonding interactions between main chain and side chain amides.

In light of their results, Pappu and coworkers argue that polyQ fibrils are not nucleated *via* a coil to β-sheet structural transition. Instead, they hypothesize that monomeric polyQ peptides adopt disordered "globule" structures that non-specifically aggregate into high molecular weight oligomers [38, 44] (Figure 1.2). According to this model, these aggregates contain interfaces that promote the formation of β-sheet-rich structures that then structurally convert into amyloid-like fibrils [38]. The Pappu model has prompted several researchers to re-evaluate the interpretation of experimental aggregation kinetics data. For example, both the Pappu [44] and Murphy groups [74] argue that the Wetzel *et al.* sedimentation assay data are insufficient to differentiate between a nucleated growth-polmerization and more complex aggregation models.

Reliable experimental data supporting the Pappu model remains scarce. Currently, only one study, by Lee *et al.* [75], claims to observe the soluble oligomeric aggregates predicted by the Pappu model. However, the presence of these oligomer aggregates may be due to an artifact that occured during sample preparation [35]. It is well known that many synthetic preparations of amyloidogenic peptides contain micro-aggregates that can heterogeneously seed fibrils [76]. To guard against this, "disaggregation" protocols [72, 76, 77] have been developed to remove these aggregates.

These protocols typically treat synthetic peptides with volatile, fluorinated solvents such as trifluoroacetic acid (TFA) and hexafluoroisopropanol (HFIP) that disrupt hydrogen bonding interactions and dissolve aggregates. As shown by Kar *et al.* [70], rigorous disaggregation of polyQ peptides do not result in the formation of high molecular weight amorphous aggre-

gates during fibril formation. Furthermore, studies [78, 79] show that HFIP can drive peptides towards oligomeric aggregate formation. Thus, the experimental evidence of oligomeric aggregates observed by Lee *et al.* [75] may be due to the failure to completely remove HFIP during disaggregation.

1.4 THE NEED FOR NEW BIOPHYSICAL TOOLS TO STUDY AMYLOID-LIKE FIBRILS

The lack of consensus regarding the structure and aggregation mechanism(s) of polyQ fibrils motivates the need to develop incisive biophysical tools that can *quantitatively* discriminate between different proposed models. The fact, for example, that three independent groups [52–55] analyzed similar X-ray fiber diffraction data and proposed four distinctly different structures highlights the difficulties associated with using conventional biophysical techniques to study polyQ amyloid-like fibrils.

Traditional methods such as X-ray diffraction and solution-state NMR typically cannot be utilized to study amyloid-like fibrils due to their insoluble and non-crytalline nature. Currently, the gold standard biophysical method to study amyloid-like fibrils is solid state NMR. Solid state NMR has been used with great success to solve several fibril structures [80–83]. An impressive array of sophisticated pulse sequences and methodologies have been developed to measure dihedral angles and distance constraints in fibrils [84–87]. In recent years, for example, the development of high-field dynamic nuclear polarization experiments have enabled the measurement of very high signal-to-noise fibril spectra [87, 88].

These sophisticated measurements, however, remain highly challenging and expensive. In order to obtain high signal-to-noise spectra, typical solid state NMR studies require extensive and costly isotopic ^{13}C and ^{15}N labeling, which can be very challenging to incorporate into proteins. Furthermore, spectral data collection typically requires long acquisition times that can take days and consume large quantities of precious samples. Finally, since spectral dispersion is often poor, many solid state NMR studies provide only qualitative information on the secondary structure of amyloid fibrils.

Ultraviolet resonance Raman (UVRR) spectroscopy is a powerful biophysical tool that can provide detailed molecular-level insights into polyQ solution-state and fibril structure. An advantage of UVRR is that, compared to NMR methods, structural information can be obtained quickly and under dilute concentrations [89]. Lednev and coworkers [90–95] have shown that UVRR is an excellent tool for probing the cross-β structure of fibrils, as well as monitoring the conformational changes that occur in amyloidgenic peptides during aggregation. In recent years, Asher and coworkers have identified several UVRR spectroscopic markers that are highly sensitive to the structure and hydrogen bonding environment of the peptide backbone [96–98], as well as amino acid side chains [99–103]. This has enabled new and deep insights into protein folding [104–107]. The primary goal of this dissertation work is to discover new spectroscopic markers that can be utilized to quantitatively investigate the structure of polyQ-rich peptides in solution-state and fibril aggregates.

2.0 UV RESONANCE RAMAN SPECTROSCOPY

2.1 CLASSICAL ELECTRODYNAMICS THEORY OF RAYLEIGH AND RAMAN SCATTERING

Figure 2.1 illustrates Rayleigh and Raman scattering from a classical electrodyanamics perspective. When electromagnetic radiation excites a molecule, its electron cloud oscillates at the frequency of the incident light, and results in a displacement of charge that induces a change in the dipole moment. Because electrons are being accelerated, energy is radiated in the form of light [108]. Most of the radiating light occurs at the same frequency (ω_I) as the incident electromagnetic radiation and is said to be elastically scattered. Assuming the wavelength of the exciting electromagnetic radiation is much larger than the dimensions of the molecule, this phenomenon is called Rayleigh scattering.

As shown in Figure 2.1, not all the light is elastically scattered. The motions of the oscillating electrons can couple to slower moving nuclear vibrational motions in the molecule. This coupling results in Raman (inelastic) scattering, where the electrons oscillate at a beat frequency. The Raman scattered light can be shifted to higher ($\omega_I + \omega_S$) or lower ($\omega_I - \omega_S$) frequencies with respect to the incident electromagnetic radiation. Scattered light that occurs at lower frequencies than the exciting incident radiation is said to be Stokes-shifted. In contrast, scattered light that occurs at higher frequencies than the incident light is anti-Stokes shifted.

Resonance Raman scattering occurs when the exciting radiation falls within the "natural" frequencies of the electron oscillators. These natural frequencies correspond to electronic transitions that, in biological molecules, typically lie in the deep ultraviolet (UV). Resonance Raman results in a tremendous enhancement in the scattering signal because the exciting

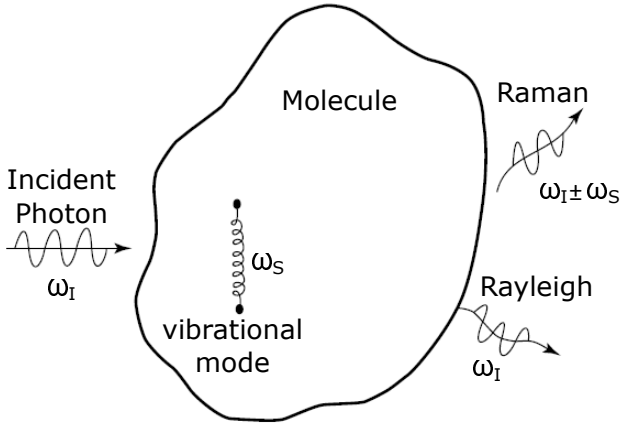


Figure 2.1: Classical physics picture of Rayleigh and Raman scattering. A molecule excited by electromagnetic radiation with a frequency of ω_I can scatter light elastically and inelastically. The elastically scattered light is known as Rayleigh scattering. The inelastically scattered light is known as Raman scattering. Adapted with permission from [109]. Copyright © (1993), American Chemical Society.

radiation is tuned into the natural frequencies of the electron oscillators. Another advantage is that only those vibrations that couple most efficiently to the electronic transition being excited are observed in resonance Raman spectra.

This second advantage confers the unique selectivity of resonance Raman spectroscopy. The vibrations of different chromophores can be selectively excited by judiciously tuning the wavelength of the excitation light so that it lies within a specific absorption band [89]. This dramatically simplifies resonance Raman spectra (Figure 2.2) and consequently relieves spectral congestion that plagues non-resonance Raman and FTIR spectra of complex biological macromolecules such as proteins. For example, in the case of myoglobin, vibrations of the heme group can be excited by tuning into the Soret band at ∼400 nm. Excitation in the UV, at the ∼220 nm, tunes into the $\pi \rightarrow \pi^*$ transition of aromatic amino acids, which enables side chain vibrations of tryptophan, tyrosine, and phenylalanine to be studied. Deeper excitation in the UV at ∼180 nm–210 nm occurs within the $\pi \rightarrow \pi^*$ transitions of amides, which enables investigations of the peptide backbone and the amino acid side chains of Gln.

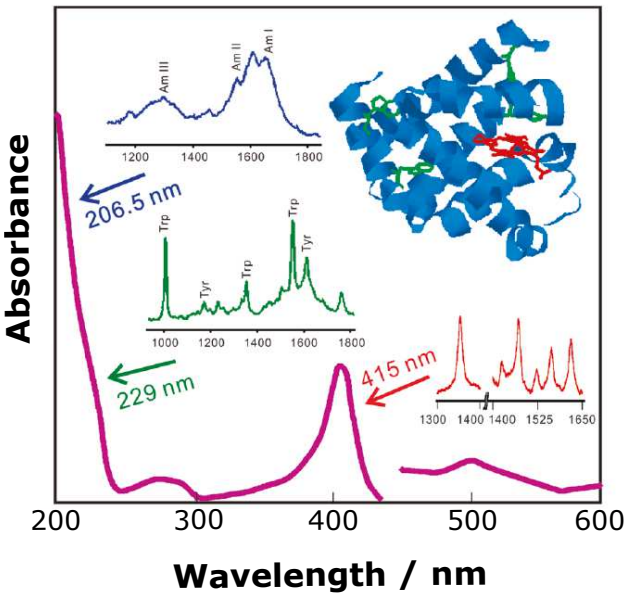


Figure 2.2: Selectivity advantage of resonance Raman spectroscopy. Excitation within different electronic absorption bands of myoglobin enable the selective enhancement of Raman bands that derive from vibrations of different chromophores. Adapted with permission from [89]. Copyright © (2012), American Chemical Society.

2.2 QUANTUM MECHANICAL THEORY OF RAYLEIGH AND RAMAN SCATTERING

The classical theory correctly predicts the existence of Rayleigh and Raman light scattering. However, it fails to provide accurate insights into the physical origins of these phenomena, as well as predict scattering intensities. The accurate theoretical treatment of Rayleigh and Raman scattering requires quantum mechanics. The traditional theoretical treatment of Rayleigh and Raman scattering begins with the dispersion equation originally derived by Kramers and Heisenberg [110]. The first quantum-mechanical derivation of the dispersion equation was given by Born, Heisenberg, and Jordan [111] using second order perturbation theory. Dirac [112] later expanded on this derivation by also quantizing the radiation field.

A key result of applying second order perturbation theory is that Rayleigh and Raman scattering belong to two photon processes, as shown in Figure 2.3. For these phenom-

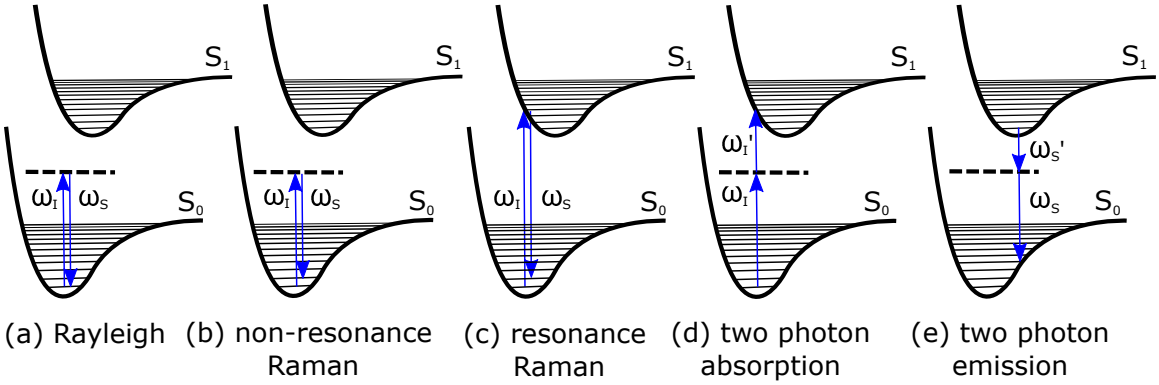


Figure 2.3: Energy diagrams showing various two photon optical processes. (a) Rayleigh scattering: a molecule absorbs a photon of frequency ω_I, which excites it into a virtual state. A photon of frequency $\omega_S = \omega_I$ is emitted, destroying the virtual state, so that the molecule's final state is the same as its initial state. (b) Non-resonance Raman scattering: a molecule absorbs a photon of frequency ω_I, which excites it into a virtual state. The virtual state is destroyed following the emission of a scattered photon of frequency $\omega_S \neq \omega_I$. (c) Resonance Raman scattering: the frequency, ω_I, of an incident photon corresponds to an electronic transition from $S_0 \rightarrow S_1$. A scattered photon of frequency ω_S is emitted so that the molecule's final state is different from the initial state. (d) Two photon absorption: a molecule absorbs two photons of frequency ω_I and ω_I'. (e) Two photon emission: a molecule emits two photons of frequency ω_S and ω_S'

ena, scattering proceeds *via* a one photon absorption and a one photon emission event. These events are coupled (Figure 2.3a–c) so that they occur nearly simultaneously [113]: the molecule in an initial state $|i\rangle$ is perturb by an incident electromagnetic field, whereupon it absorbs a photon of frequency ω_I and is excited into an intermediate, so-called "virtual" state. The molecule then emits a "scattered" photon of frequency ω_S, which destroys this virtual state, and brings it to a final state $|f\rangle$.

2.2.1 The Kramers-Heisenberg Dispersion relation

The central equation to understand Rayleigh and Raman scattering is the Kramers-Heisenberg dispersion relation, which we will derive here. To derive the dispersion relation,

we begin with the transition rate (see APPENDIX A, Derivation of the Transition Scattering Rate for a detailed derivation and discussion) associated with two photon scattering [114, 115]:

$$W_{fi} = \frac{\pi E_I^2 E_S^2}{8\hbar^3} \sum_r \left| \frac{\langle f|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|r\rangle \langle r|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|r\rangle \langle r|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right|^2 \rho_f \tag{2.1}$$

where E, ϵ, ω are the amplitudes, polarization directions, and frequencies, respectively, of the incident (denoted by the subscript I) and scattered (denoted by the subscript S) electromagnetic fields. The term, ω_{ri}, is equal to the frequency difference between an arbitrary eigenstate, $|r\rangle$, and the initial state of the system, $|i\rangle$ (*i.e.* $\omega_r - \omega_i$). The transition dipole operators, $\hat{\mu}$, couple the molecule to the incident electromagnetic field perturbing the system, and the subscripts, σ and ρ, denote the directions of the transition moments. The density of $|f\rangle$ states is denoted by ρ_f.

The differential scattering cross section, $d\sigma/d\Omega_f$, can be derived from the scattering rate by dividing eq. 2.1 by the incident photon flux, Φ (eq. 2.2), and substituting E_S and ρ_f with eqs. 2.3 and 2.4 [115] (see eq. A.15):

$$\Phi = \frac{E_I^2 c}{8\pi\hbar\omega_I} \tag{2.2}$$

$$E_S = \frac{8\pi\hbar\omega_S}{V} \tag{2.3}$$

$$\rho_f = \frac{\omega_S^2 V}{(2\pi c)^3 \hbar} d\Omega_f \tag{2.4}$$

$$\frac{d\sigma}{d\Omega_f} = \frac{\omega_I \omega_S^3}{c^4} \frac{1}{\hbar^2} \sum_r \left| \frac{\langle f|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|r\rangle \langle r|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|r\rangle \langle r|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right|^2 \tag{2.5}$$

where c is the speed of light and $d\Omega_f$ is the scattering solid angle.

Eq. 2.5 is the famous Kramers-Heisenberg dispersion relation. The scattering cross section, which is proportional to the scattering intensity, can be derived from this equation

by averaging over all molecular orientations, integrating over the full 4π solid angle, and summing over all incident and scattered light polarizations [115]:

$$\sigma = \frac{8\pi e^4 \omega_I \omega_S^3}{9c^4} \frac{1}{\hbar^2} \sum_{\rho,\sigma} \sum_{r} \left| \frac{\langle f|\hat{r}_\rho|r\rangle \langle r|\hat{r}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{r}_\sigma|r\rangle \langle r|\hat{r}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right|^2 \tag{2.6}$$

where we utilize the fact that $\hat{\mu} = e\hat{r}$ ($\hat{r}$ is the position operator and e is the charge of an electron). The cross section equation can be further simplified by introducing the second rank tensor, $(\alpha_{\rho\sigma})_{fi}$, which is known as the molecular polarizability:

$$(\alpha_{\rho\sigma})_{fi} = \frac{1}{\hbar} \sum_{r} \left[\frac{\langle f|\hat{r}_\rho|r\rangle \langle r|\hat{r}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{r}_\sigma|r\rangle \langle r|\hat{r}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right] \tag{2.7}$$

so that eq. 2.6 now becomes:

$$\sigma = \frac{8\pi e^4 \omega_I \omega_S^3}{9c^4} \sum_{\rho,\sigma} \left| (\alpha_{\rho\sigma})_{fi} \right|^2 \tag{2.8}$$

The quantum mechanical origins of Rayleigh and Raman scattering, as discussed earlier in regards to Figure 2.3, is revealed in eq. 2.7 and 2.8. These expressions indicate that the virtual state is actually a linear combination of all the eigenstates in the system. Since the molecule does not possess an eigenstate that is resonant with this virtual state, energy conservation appears to be violated. However, this is not the case since the absorption and emission events occur nearly simultaneously. Given the short lifetime of the intermediate state, Heisenberg's uncertainty principle ($\Delta E \Delta t \geq \hbar$) states that the energy uncertainty (ΔE) of the system is very large over the short time interval (Δt) during which the virtual state exists. Thus, any $|r\rangle$ eigenstate of the system can be momentarily produced, thereby maintaining energy conservation [113, 114].

For Rayleigh scattering, there is no frequency shift in the scattered light relative to the incident light ($\omega_I = \omega_S$), so that the initial and final states of the excited molecule are the same. From eq. 2.8, it can be seen that, under this condition, the cross section scales as ω_I^4, in accordance Rayleigh's famous scattering law. This ω_I^4 frequency dependence explains, for example, the color of the sky. Molecules in the atmosphere scatter shorter (blue) wavelength light more efficiently than longer (red) wavelength light. This is why the color of the sky is blue during the day.

In Raman scattering, there is an exchange of a quantum of energy so that the initial and final states of the molecule are different. As shown in Figure 2.3b, the frequency difference between the ω_I and ω_S corresponds to a vibrational transition from state $|i\rangle$ to $|f\rangle$. In this sense, Raman spectroscopy is complementary to infrared (IR) spectroscopy. However, because Raman is a two photon scattering process and IR is a one photon absorption process, the selection rules between these two spectroscopies are different.

2.3 NON-RESONANCE VS. RESONANCE SCATTERING

The traditional approach to understand resonance Raman scattering (Figure 2.3c) is to utilize the vibronic theory championed primarily by Albrecht and coworkers [115–118]. This theory treats Rayleigh and Raman scattering in the framework of vibronic spectroscopies, where, the resonance scattering intensities of the electronic ground state vibrations are intimately tied to transitions of excited vibronic states. In this section, we examine the vibronic theory by Albrecht in order to elucidate important physical aspects of resonance and non-resonance Raman scattering.

2.3.1 General Vibronic Theory

The wave functions in eq. 2.7 depend on *both* the electronic and vibrational states. Albrecht [117] first assumed that the electronic and vibrational state wavefunctions are separable by invoking the Born-Oppenheimer approximation:

$$|\Psi_m\rangle = |\phi_e(q,Q)\chi_{em}(Q)\rangle = |\phi_e(q,Q)\rangle\,|\chi_{em}(Q)\rangle = |\phi_e\rangle\,|\chi_{em}\rangle \tag{2.9}$$

where $|\phi_e\rangle$ is the wave function for an electronic state, e, which depends on the complete set of internal coordinates of both the nuclei (Q) and the electrons (q). The state $|\chi_{em}\rangle$ is the vibrational wave function associated with the resonant electronic excited state, e, and the vibrational state, m. Inserting eq. 2.9 into eq. 2.7 gives the following equation:

$$(\alpha_{\rho\sigma})_{gf',gi'} = \frac{1}{\hbar}\sum_{e}\sum_{r'}\left[\frac{\langle\chi_{gf'}|\langle\phi_g|\hat{r}_\rho|\phi_e\rangle|\chi_{er'}\rangle\langle\chi_{er'}|\langle\phi_e|\hat{r}_\sigma|\phi_g\rangle|\chi_{gi'}\rangle}{\omega_{er'}-\omega_{gi'}-\omega_I} + \frac{\langle\chi_{gf'}|\langle\phi_g|\hat{r}_\sigma|\phi_e\rangle|\chi_{er'}\rangle\langle\chi_{er'}|\langle\phi_e|\hat{r}_\rho|\phi_g\rangle|\chi_{gi'}\rangle}{\omega_{er'}-\omega_{gi'}+\omega_S}\right] \tag{2.10}$$

where $\omega_{ri} = \omega_{er'} - \omega_{gi'}$. The electronic wave functions can be expanded using a Taylor's series about the nuclear coordinates around the equilibrium ground state structure. To first order theory, the result of this Herzberg-Teller expansion [117] is:

$$|\phi_g\rangle = |\phi_g^{(0)}\rangle + \frac{1}{\hbar}\sum_{a}\sum_{t\neq g}\frac{(h_a^{(0)})_{es}\cdot\hat{Q}_a}{\omega_g^{(0)}-\omega_t^{(0)}}|\phi_t^{(0)}\rangle \tag{2.11}$$

$$|\phi_e\rangle = |\phi_e^{(0)}\rangle + \frac{1}{\hbar}\sum_{a}\sum_{s\neq e}\frac{(h_a^{(0)})_{gt}\cdot\hat{Q}_a}{\omega_g^{(0)}-\omega_s^{(0)}}|\phi_s^{(0)}\rangle \tag{2.12}$$

where $|\phi_g^{(0)}\rangle$, $|\phi_e^{(0)}\rangle$, $|\phi_s^{(0)}\rangle$, and $|\phi_t^{(0)}\rangle$ are the unperturbed electronic wave functions for states g, e, s, and t, respectively. State g is the electronic ground state, while s and t are non-resonant electronic excited states. The operator, $\hat{Q}_a$, is the displacement operator for the ath normal mode. The Herzberg-Teller coupling integrals, $(h_a^{(0)})_{gt}$ and $(h_a^{(0)})_{es}$, are defined as:

$$(h_a^{(0)})_{gt} = \langle\phi_g^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_t^{(0)}\rangle \tag{2.13}$$

$$(h_a^{(0)})_{es} = \langle\phi_s^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_s^{(0)}\rangle \tag{2.14}$$

Substituting eqs. 2.11 and 2.12 into eq. 2.10 gives the following equation [115] (see APPENDIX A, Derivation of the Albrecht A, B, and C Terms for details):

$$(\alpha_{\rho\sigma})_{gf',gi'} \approx A + B + C \tag{2.15}$$

where,

$$A = \frac{1}{\hbar}\sum_{e\neq g}\sum_{r'}\left[\frac{\langle\phi_g|\hat{r}_\rho|\phi_e\rangle\langle\phi_e|\hat{r}_\sigma|\phi_g\rangle}{\omega_{er'}-\omega_{gi'}-\omega_I} + \frac{\langle\phi_g|\hat{r}_\sigma|\phi_e\rangle\langle\phi_e|\hat{r}_\rho|\phi_g\rangle}{\omega_{er'}-\omega_{gi'}+\omega_S}\right]\langle\chi_{gf'}|\chi_{er'}\rangle\langle\chi_{er'}|\chi_{gi'}\rangle \tag{2.16}$$

$$B = \frac{1}{\hbar^2} \sum_{e \neq g} \sum_{r'} \sum_{a} \sum_{s \neq e} \Bigg\{ \Bigg[\frac{\langle \phi_g^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle}{\omega_{er'} - \omega_{gi'} - \omega_I}$$
$$+ \frac{\langle \phi_g^{(0)} | \hat{r}_\sigma | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | \hat{r}_\rho | \phi_g^{(0)} \rangle}{\omega_{er'} - \omega_{gi'} + \omega_S} \Bigg] \Bigg[\frac{\langle \chi_{gf'} | \chi_{er'} \rangle \langle \chi_{er'} | \hat{Q}_a | \chi_{gi'} \rangle}{\omega_e^{(0)} - \omega_s^{(0)}} \Bigg]$$
$$+ \Bigg[\frac{\langle \phi_g^{(0)} | \hat{r}_\rho | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle}{\omega_{er'} - \omega_{gi'} - \omega_I}$$
$$+ \frac{\langle \phi_g^{(0)} | \hat{r}_\sigma | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\rho | \phi_g^{(0)} \rangle}{\omega_{er'} - \omega_{gi'} + \omega_S} \Bigg] \Bigg[\frac{\langle \chi_{gf'} | \hat{Q}_a | \chi_{er'} \rangle \langle \chi_{er'} | \chi_{gi'} \rangle}{\omega_e^{(0)} - \omega_s^{(0)}} \Bigg] \Bigg\} \quad (2.17)$$

$$C = \frac{1}{\hbar^2} \sum_{e \neq g} \sum_{r'} \sum_{a} \sum_{t \neq e} \Bigg\{ \Bigg[\frac{\langle \phi_g^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_t^{(0)} \rangle \langle \phi_t^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_g^{(0)} \rangle}{\omega_{er'} - \omega_{gi'} - \omega_I}$$
$$+ \frac{\langle \phi_g^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_t^{(0)} \rangle \langle \phi_t^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_g^{(0)} \rangle}{\omega_{er'} - \omega_{gi'} + \omega_S} \Bigg] \Bigg[\frac{\langle \chi_{gf'} | \chi_{er'} \rangle \langle \chi_{er'} | \hat{Q}_a | \chi_{gi'} \rangle}{\omega_e^{(0)} - \omega_s^{(0)}} \Bigg]$$
$$+ \Bigg[\frac{\langle \phi_g^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_t^{(0)} \rangle \langle \phi_t^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle}{\omega_{er'} - \omega_{gi'} - \omega_I}$$
$$+ [\frac{\langle \phi_g^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_t^{(0)} \rangle \langle \phi_t^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle}{\omega_{er'} - \omega_{gi'} + \omega_S} \Bigg] \Bigg[\frac{\langle \chi_{gf'} | \hat{Q}_a | \chi_{er'} \rangle \langle \chi_{er'} | \chi_{gi'} \rangle}{\omega_g^{(0)} - \omega_t^{(0)}} \Bigg] \Bigg\} \quad (2.18)$$

For a particular vibration, a, the scattering intensities that derive from the B and C terms due to Herzberg-Teller coupling are proportional to the displacement, Q_a, and inversely proportional to the frequency differences, $\omega_g^{(0)} - \omega_t^{(0)}$ and $\omega_e^{(0)} - \omega_s^{(0)}$. The frequency dependence of the B and C terms means that significant coupling will occur only when the $|s\rangle$ and $|t\rangle$ electronic states are close in energy to the electronic ground state, $|g\rangle$, and the resonant electronic excited state, $|e\rangle$. Typically, the energy difference between $|g\rangle$ and $|t\rangle$ is much larger than that between $|e\rangle$ and $|s\rangle$ (*i.e.*, $\omega_g^{(0)} - \omega_t^{(0)} \gg \omega_e^{(0)} - \omega_s^{(0)}$). This means that, in typical cases, the C term can be neglected, so that eq. 2.15 becomes [115]:

$$(\alpha_{\rho\sigma})_{gf',gi'} \approx A + B \quad (2.19)$$

2.3.2 Non-Resonance Rayleigh and Raman Scattering

Eq. 2.19 can used to understand non-resonance and resonance scattering. We will first briefly consider the non-resonance case, wherein the frequency of the exciting electric field is far away from an electronic absorption band such that $|\omega_{er'} - \omega_{gi'} - \omega_I| \gg 0$. In this situation, $\omega_{er'} - \omega_{gi'} - \omega_I \approx \omega_{er'} - \omega_{gi'} + \omega_S$ and $\omega_{er'} - \omega_{gi'} \approx \omega_e^{(0)} - \omega_g^{(0)}$. Using these approximations, eq. 2.19 can be written as:

$$A = \frac{1}{\hbar} \sum_{e \neq g} \left[\frac{\langle \phi_g | \hat{r}_\rho | \phi_e \rangle \langle \phi_e | \hat{r}_\sigma | \phi_g \rangle}{\omega_e^{(0)} - \omega_g^{(0)} - \omega_I} + \frac{\langle \phi_g | \hat{r}_\sigma | \phi_e \rangle \langle \phi_e | \hat{r}_\rho | \phi_g \rangle}{\omega_e^{(0)} - \omega_g^{(0)} + \omega_I} \right] \langle \chi_{gf'} | \chi_{gi'} \rangle \tag{2.20}$$

$$\begin{aligned} B = \frac{1}{\hbar^2} \sum_{e \neq g} \sum_a \sum_{s \neq e} \Bigg\{ & \\ & \Bigg[\frac{\langle \phi_g^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle + \langle \phi_g^{(0)} | \hat{r}_\rho | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle}{\omega_e^{(0)} - \omega_g^{(0)} - \omega_I} \\ & + \frac{\langle \phi_g^{(0)} | \hat{r}_\sigma | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | \hat{r}_\rho | \phi_g^{(0)} \rangle + \langle \phi_g^{(0)} | \hat{r}_\sigma | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\rho | \phi_g^{(0)} \rangle}{\omega_e^{(0)} - \omega_g^{(0)} + \omega_I} \Bigg] \\ & \times \left[\frac{\langle \chi_{gf'} | \hat{Q}_a | \chi_{gi'} \rangle}{\omega_e^{(0)} - \omega_s^{(0)}} \right] \Bigg\} \end{aligned} \tag{2.21}$$

To derive these equations, all r' vibrational states were summed over using the identity shown in A.9. Eqs. 2.20 and 2.21 can be further simplified by assuming that the wavefunctions are all real, so that the numerators belonging to the $\omega_e^{(0)} - \omega_g^{(0)} - \omega_I$ and $\omega_e^{(0)} - \omega_g^{(0)} + \omega_I$ terms are equal. Thus, using the following algebraic identity,

$$\frac{1}{a-b} + \frac{1}{a+b} = \frac{2a}{a^2 - b^2} \tag{2.22}$$

eqs. 2.20 and 2.21 can be simplified to:

$$A = \frac{1}{\hbar} \sum_{e \neq g} \left[\frac{2(\omega_e^{(0)} - \omega_g^{(0)})}{(\omega_e^{(0)} - \omega_g^{(0)})^2 - \omega_I^2} \right] \langle \phi_g | \hat{r}_\rho | \phi_e \rangle \langle \phi_e | \hat{r}_\sigma | \phi_g \rangle \langle \chi_{gf'} | \chi_{gi'} \rangle \tag{2.23}$$

$$\begin{aligned} B = \frac{1}{\hbar^2} \sum_{e \neq g} \sum_a \sum_{s \neq e} \Bigg\{ & \left[\frac{2(\omega_e^{(0)} - \omega_g^{(0)})}{(\omega_e^{(0)} - \omega_g^{(0)})^2 - \omega_I^2} \right] \left[\frac{\langle \chi_{gf'} | \hat{Q}_a | \chi_{gi'} \rangle}{\omega_e^{(0)} - \omega_s^{(0)}} \right] \\ & \times \langle \phi_g^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle \\ & + \langle \phi_g^{(0)} | \hat{r}_\rho | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle \Bigg\} \end{aligned} \tag{2.24}$$

It can be seen that the A term is non-zero only when $|\chi_{gi'}\rangle = |\chi_{gf'}\rangle$. Therefore, the A term is solely responsible for Rayleigh scattering in non-resonance cases. In contrast, B term is non-zero on three conditions: $|\chi_{gf'}\rangle = |\chi_{gi'} \pm 1\rangle$; the Herzberg-Teller coupling integrals are non-zero; and there are no interference effects that occur when summing over the electronic states. Thus, according to Albrecht's vibronic theory, the B term is solely responsible for non-resonance Raman scattering intensities of fundamentals.

2.3.3 Resonance Rayleigh and Raman Scattering

We now consider the case of resonance, where the excitation frequency lies within a molecule's electronic absorption band. As ω_I approaches $\omega_{er'} - \omega_{gi'}$, eq. 2.10 is no longer valid because the scattering amplitude that derives from the term containing $\omega_{er'} - \omega_{gi'} - \omega_I$ becomes infinite. In reality, radiation dampening occurs due to the finite lifetime of populating the eigenstates. To account for this, we introduce a phenomenological damping coefficient, Γ, so that eq. 2.10 can be re-written as:

$$(\alpha_{\rho\sigma})_{gf',gi'} = \frac{1}{\hbar} \sum_{e} \sum_{r'} \left[\frac{\langle \chi_{gf'}| \langle \phi_g|\hat{r}_\rho|\phi_e\rangle |\chi_{er'}\rangle \langle \chi_{er'}| \langle \phi_e|\hat{r}_\sigma|\phi_g\rangle |\chi_{gi'}\rangle}{\omega_{er'} - \omega_{gi'} - \omega_I + i\Gamma} + \frac{\langle \chi_{gf'}| \langle \phi_g|\hat{r}_\sigma|\phi_e\rangle |\chi_{er'}\rangle \langle \chi_{er'}| \langle \phi_e|\hat{r}_\rho|\phi_g\rangle |\chi_{gi'}\rangle}{\omega_{er'} - \omega_{gi'} + \omega_S + i\Gamma} \right] \tag{2.25}$$

Using this expression, eq. 2.19 can be modified in two different ways. First, the summation over e disappears, since the molecule is being excited into a specific electronic excited state. Second, the "non-resonant" $\omega_{er'} - \omega_{gi'} + \omega_S$ term in eq. 2.25 can be neglected because it is significantly smaller than the $\omega_{er'} - \omega_{gi'} - \omega_I$ term. Thus, eqs. 2.16 and 2.17 can be re-written as:

$$A = \frac{1}{\hbar} \langle \phi_g|\hat{r}_\rho|\phi_e\rangle \langle \phi_e|\hat{r}_\sigma|\phi_g\rangle \sum_{r'} \left[\frac{\langle \chi_{gf'}|\chi_{er'}\rangle \langle \chi_{er'}|\chi_{gi'}\rangle}{\omega_{er'} - \omega_{gi'} - \omega_I + i\Gamma} \right] \tag{2.26}$$

$$B = \frac{1}{\hbar^2} \sum_{r'} \sum_{a} \sum_{s \neq e} \left\{ \left[\frac{\langle \phi_g^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle}{\omega_{er'} - \omega_{gi'} - \omega_I + i\Gamma} \right] \left[\frac{\langle \chi_{gf'} | \chi_{er'} \rangle \langle \chi_{er'} | \hat{Q}_a | \chi_{gi'} \rangle}{\omega_e^{(0)} - \omega_s^{(0)}} \right] \right.$$
$$\left. + \left[\frac{\langle \phi_g^{(0)} | \hat{r}_\rho | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle}{\omega_{er'} - \omega_{gi'} - \omega_I + i\Gamma} \right] \left[\frac{\langle \chi_{gf'} | \hat{Q}_a | \chi_{er'} \rangle \langle \chi_{er'} | \chi_{gi'} \rangle}{\omega_e^{(0)} - \omega_s^{(0)}} \right] \right\} \tag{2.27}$$

It is impossible to sum over all $|r'\rangle$ vibrational states, as was done in deriving eqs. 2.23 and 2.24. This is because in resonance, the frequency terms in the denominators of the summations are now sensitive to only a subset of $|r'\rangle$ states.

A comparison of these equations with 2.23 and 2.24 reveals interesting differences in the mechanism of Rayleigh and Raman scattering between the resonance and non-resonance cases. Unlike the non-resonance situation, the resonance A and B terms are both responsible for Rayleigh and Raman scattering processes. In the case of the A term, Albrecht [117] notes that this is because the vibrational wavefunctions of the electronic ground and resonant excited states are generally not solutions to the same Schrödinger equation. Therefore, the Franck-Condon factors ($\langle \chi_{gf'} | \chi_{er'} \rangle \langle \chi_{er'} | \chi_{gi'} \rangle$) in the A term need not be zero.

It is also interesting to note that the contribution of the A term to the Raman scattering intensity of a vibrational mode does not depend on vibronic mixing between the resonant excited state, e, with other electronic states. In contrast, B term scattering depends on vibronic mixing of the resonant state, e, with other electronic states. The $1/(\omega_e^{(0)} - \omega_s^{(0)})$ dependence of the terms containing the Herzberg-Teller integrals in eq. 2.27 means that the lowest lying e to s state electronic transitions will be responsible for most of the B term scattering intensity.

2.3.4 Resonance Raman and the Electronic Excited State Geometry

Eqs. 2.26 and 2.27 indicate that the resonance Raman intensities are inherently sensitive to details of the electronic excited state potential energy surface. This can be most easily seen in the Albrecht A term, where only one electronic transition contributes to the resonance Raman scattering intensity. In the A term, information about the excited state geometry is encoded in the Franck-Condon factors that are in the summation of eq. 2.26. The electronic

excited state must be displaced along a particular coordinate for at least one of the Franck-Condon factors to be non-zero.

The simplest approach to determine the Franck-Condon integrals is to assume that the electronic ground and excited states are both harmonic and differ only in their equilibrium positions [119]. In this limit, no Duschinsky rotation occurs [119], and the vibrational frequencies of the normal modes do not change between the ground and excited states. Each vibrational j^{th} mode can be treated as a pair of ground and excited state harmonic oscillators that are separated in energy by some frequency, ω_j, and displaced relative to each other along a particular coordinate by an origin shift, Δ. As a result, the multidimensional Franck-Condon factors can be written as the products of one dimensional overlap integrals [119]. *i.e.*,

$$\langle \chi_{gf'} | \chi_{er'} \rangle = \prod_{j=1}^{N} \langle \chi_{gf'_j} | \chi_{er'_j} \rangle \tag{2.28}$$

Using eq. 2.28, the resonance Raman polarizability corresponding to the transition from $|\chi_{gf'} = 0_1\rangle$ to $|\chi_{gf'} = 1_1\rangle$ for the $j = 1$ Raman mode (denoted by the subscript 1) is [119]:

$$\begin{aligned} \alpha_{0\to1} = & \frac{1}{\hbar} \langle \phi_g | \hat{r}_\rho | \phi_e \rangle \langle \phi_e | \hat{r}_\sigma | \phi_g \rangle \\ & \times \sum_{m_1} \sum_{m_2} \cdots \sum_{m_{3N-6}} \left[\frac{\langle 1_1 | m_1 \rangle \langle m_1 | 0_1 \rangle \prod_{j=2}^{3N-6} \langle 0_j | m_j \rangle \langle m_j | 0_j \rangle}{\omega_e^0 - \omega_g^0 + \sum_{j=1}^{3N-6} (m_j)\omega_j - \omega_I + i\Gamma} \right] \end{aligned} \tag{2.29}$$

where $|\chi_{er'_j}\rangle = |m_j\rangle$.

The one dimensional Franck-Condon factors can be re-written using a set of recursion equations derived by Manneback [120]. Using these recursion relations, the Frank-Condon factors can be re-written as [119]:

$$\langle 0_j | m_j \rangle \langle m_j | 0_j \rangle = \frac{\Delta_j^{2m}}{2^m m!} e^{-\Delta_j^2/2} \tag{2.30}$$

$$\langle 1_j | m_j \rangle \langle m_j | 0_1 \rangle = \frac{\Delta_j}{\sqrt{2}} [\langle 0_1 | m_1 \rangle \langle m_1 | 0_1 \rangle - \langle 0_1 | m_1 - 1 \rangle \langle m_1 - 1 | 0_1 \rangle] \tag{2.31}$$

in the limit that the vibrational frequencies for the electronic ground and the excited state normal modes do not change. Using these equations, 2.29 can be written as [119]:

$$\alpha_{0\to 1} = \left(\frac{1}{\hbar}\right)\left(\frac{\Delta_1}{\sqrt{2}}\right)\langle\phi_g|\hat{r}_\rho|\phi_e\rangle\langle\phi_e|\hat{r}_\sigma|\phi_g\rangle \times \sum_{m_1}\sum_{m_2}\cdots\sum_{m_{3N-6}}\left[\frac{(\langle 0_1|m_1\rangle\langle m_1|0_1\rangle - \langle 0_1|m_1-1\rangle\langle m_1-1|0_1\rangle)\prod_{j=2}^{3N-6}\langle 0_j|m_j\rangle\langle m_j|0_j\rangle}{\omega_e^0 - \omega_g^0 + \sum_{j=1}^{3N-6}(m_j)\omega_j - \omega_I + i\Gamma}\right] \tag{2.32}$$

Eq. 2.32 shows that the resonance Raman cross sections for fundamentals are proportional to Δ^2, the square of the displacement between the electronic ground and excited states along a particular normal coordinate. Thus, eq. 2.32 shows that resonance Raman band intensities are sensitive to the geometry changes that occur between the electronic ground and excited states of molecules.

3.0 UV RESONANCE RAMAN INVESTIGATION OF THE AQUEOUS SOLVATION DEPENDENCE OF PRIMARY AMIDE VIBRATIONS

Adapted with permission from: David Punihaole, Ryan S. Jakubek, Elizabeth M. Dahlburg, Zhenmin Hong, Nataliya S. Myshakina, Steven Geib, Sanford A. Asher. UV Resonance Raman Investigation of the Aqueous Solvation Dependence of Primary Amide Vibrations. *The Journal of Physical Chemistry B*, **2015**, 119, 3931–3939.

Author Contributions: D.P. acquired and analyzed the UVRR data with the assistance of R.S.J. and E.M.D. D.P. and S.G. acquired and analyzed X-ray diffraction data. Z.H. performed DFT calculations and normal mode analysis. D.P. and N.S. assisted in analyzing the results from the DFT calculations. The manuscript was prepared by D.P. and S.A.A. with the assistance of R.S.J. and E.M.D.

We investigated the normal mode composition and the aqueous solvation dependence of the primary amide vibrations of propanamide. Infrared, normal Raman, and UVRR spectroscopy were applied in conjunction with density functional theory (DFT) to assign the vibrations of crystalline propanamide. We examined the aqueous solvation dependence of the primary amide UVRR bands by measuring spectra in different acetonitrile/water mixtures. As previously observed in the UVRR spectra of N-methylacetamide, all of the resonance enhanced primary amide bands, except for the Amide I (Am I), show increased UVRR cross sections as the solvent becomes water-rich. These spectral trends are rationalized by a

model wherein the hydrogen bonding and the high dielectric constant of water stabilizes the *ground state* dipolar $^{-}O-C=NH_2^{+}$ resonance structure over the neutral $O=C-NH_2$ resonance structure. Thus, vibrations with large C−N stretching show increased UVRR cross sections because the C−N displacement between the electronic ground and excited state increases along the C−N bond. In contrast, vibrations dominated by C=O stretching, such as the Am I, show a decreased displacement between the electronic ground and excited state, which result in a decreased UVRR cross section upon aqueous solvation. The UVRR primary amide vibrations can be used as sensitive spectroscopic markers to study the local dielectric constant and hydrogen bonding environments of the primary amide side chains of Gln and Asn.

3.1 INTRODUCTION

The primary amide functional group is of significant biological interest since it is found in the side chains of Gln and Asn. These side chains may be of structural and functional significance to peptides and proteins since they can participate in both intra- and intermolecular hydrogen bonding, which may be important to the formation and stabilization of prion and amyloid-like fibril aggregates that are involved in a number of protein diseases [53, 54, 56, 121, 122]. Given the importance of these primary amide groups, it is of great value to find spectroscopic markers that can be used to monitor the Asn and Gln hydrogen bonding and dielectric environments.

UVRR spectroscopy is a powerful tool for studying the conformations of proteins [89], as well as for determining hydrogen bonding, protonation states, and local dielectric environments of aromatic amino acids [101], arginine [103], and histidine [123, 124] side chains. Deep UV excitation (∼200 nm) selectively enhances the peptide bond secondary amide vibrations of the protein backbone [89] and the primary amide vibrations of the Asn and Gln side chains [125]. Investigations of the secondary amide vibrations [96, 97, 105, 106, 126–128] have developed a deep understanding of the spectral dependence of the peptide bond secondary structure and its hydrogen bonding. This understanding has enabled incisive investigations

of protein and peptide structure. The work presented here is developing a similar deep understanding of the structural, hydrogen bonding, and dielectric environmental dependence of primary amide vibrations.

We investigated propanamide, one of the simplest primary amide compounds with a structure similar to that of the Asn and Gln side chains. We assigned the vibrations observed in the infrared, non-resonance Raman, and UVRR spectra of crystalline propanamide with the aid of DFT. We then examined the solution behavior of the primary amide UVRR bands in mixtures of acetonitrile and water. The primary amide bands are very sensitive to their hydrogen bonding and dielectric environments. These bands will be useful as spectroscopic probes to monitor the side chain environment and structure of Gln and Asn.

3.2 EXPERIMENTAL SECTION

3.2.1 Materials

Propanamide ($CH_3CH_2CONH_2$, 97% purity) and acetonitrile (HPLC, far-UV grade) were purchased from Acros Organics. N-methylacetamide ($CH_3CH_2CONHCH_3$, NMA, $\geq$99% purity) and sodium perchlorate ($NaClO_4$, $\geq$98% purity) were purchased from Sigma-Aldrich. D_2O (99.9% atom D purity) was purchased from Cambridge Isotope Laboratories, Inc.

3.2.2 Sample Preparation

The propanamide solid samples consisted of a crystalline powder, which was used without further purification or re-crystallization. N-deuterated propanamide crystals were prepared by multiple re-crystallizations in D_2O. Propanamide solutions were prepared in H_2O at 10 mM concentration. Samples prepared in mixtures of acetonitrile and water were prepared at 30 mM concentrations. For these experiments, $NaClO_4$ (200 mM) was used as an internal standard. The N-deuterated propanamide solution was prepared at 30 mM concentration in pure D_2O.

3.2.3 Infrared and Non-Resonance Raman Spectroscopy

The mid-infrared spectrum of propanamide crystals was measured in the 600–4000 cm^{-1} region at 1 cm^{-1} resolution. The data were collected using a Perkin-Elmer model Spectrum 100 series FTIR equipped with a Universal diamond ATR. The propanamide crystals were lightly ground for ∼30 s using a mortar and pestle in order to ensure good optical contact between the sample and the diamond crystal. The sample was placed on the diamond crystal and a force of ∼145 N was applied using a pressure arm.

The visible excitation Raman spectra of propanamide crystals were measured using a Renishaw inVia Raman spectrometer equipped with a research-grade Leica microscope. Spectra were collected using a 5× objective lens with a ∼2 cm^{-1} resolution spectrometer. The 633 nm exciting line was generated by a HeNe laser. The 380 cm^{-1}, 918 cm^{-1}, 1376 cm^{-1}, 2249 cm^{-1}, and 2942 cm^{-1} bands of acetonitrile [129] were used for calibration.

3.2.4 UVRR Spectroscopy

The UVRR spectra of crystalline propanamide were measured using cw 229 nm light generated by an Innova 300 FreD frequency doubled Ar^+ laser [130]. Solid samples were spun by using a cylindrical brass rotation cell to prevent thermal degradation or photodegradation. A SPEX Triplemate spectrograph, modified for use in the UV, was utilized to disperse the Raman scattered light. A Spec-10 system charge-coupled device (CCD) camera (Princeton Instruments, Model 735-0001) was employed to detect the scattered light.

UVRR solution-state measurements were made using ∼204 nm excitation. The UV light was generated by Raman shifting the third harmonic of an Nd:YAG Infinity laser (Coherent, Inc.) in H_2 gas (∼30 psi) and selecting the fifth anti-Stokes line. Solutions were circulated using a thermostatted (20 °C) flow cell [104] to prevent accumulation of photodegradation products. The scattered light was dispersed and imaged using a double monochromator, modified for use in the UV in a subtractive configuration, [131] and detected with a Spec-10 CCD camera.

3.2.5 UV Absorption Measurements

Absorption spectra were taken of 30 mM propanamide and 10 mM N-methylacetamide solutions dissolved in acetonitrile and water. A Varian Cary 5000 spectrophotometer with a 0.2 mm path length quartz cuvette was used for all measurements.

3.2.6 Raman Cross Section Calculations

The UVRR scattering cross sections of the spectrally deconvoluted propanamide Raman bands (see APPENDIX B for details) were calculated using the following equation [125]:

$$\sigma_i = \frac{I_i k_r C_r \sigma_r}{I_r k_i C_i} \left(\frac{\epsilon_i + \epsilon_{ex}}{\epsilon_r + \epsilon_{ex}} \right) \tag{3.1}$$

where σ_i is the cross section of the ith propanamide Raman band, and σ_r is the cross section of the 932 cm^{-1} ClO_4^- stretching band from our internal standard. C_r and C_i are the concentrations of $NaClO_4$ and propanamide, respectively. The factors k_r and k_i are the spectrometer efficiencies at the 932 cm^{-1} and ith propanamide Raman bands. Assuming only the analyte absorbs, the factors ϵ_i, ϵ_r, and ϵ_{ex} are the molar absorptivities at the ith propanamide Raman band, the 932 cm^{-1} band, and the excitation wavelength, respectively. The expression in the parenthesis corrects the cross section measurement for self-absorption [132, 133]. The Raman cross section of the 932 cm^{-1} reference band, σ_r, at 204 nm excitation, was estimated to be $\sim 1.18 \times 10^{-27}$ cm^2·molecule^{-1}·sr^{-1} by extrapolating the Raman cross section measurements of Dudik *et al.* [134]

3.3 COMPUTATIONAL SECTION

The DFT calculations [135] were performed using the GAUSSIAN 09 package [136] with the M06-2X functional [137] and the 6-311++g** basis set. The calculated frequencies of the propanamide vibrations were calculated using a harmonic approximation and scaled linearly to the average band frequencies observed experimentally. The calculations simulated solvation implicitly by placing the propanamide molecule in an ellipsoidal cavity surrounded

by a polarizable continuum dielectric modeled to simulate water. The potential energy distribution (PED) of each vibration was obtained from the GAUSSIAN output files by employing the GAR2PED program [138].

3.4 RESULTS AND DISCUSSION

3.4.1 Normal Mode Analysis

The infrared and Raman spectra of propanamide were reported previously [139–144]. The most detailed normal mode analysis and assignments were carried out by Kuroda *et al.* [143] and Nandini and Sathyanarayana [144]. Kuroda *et al.* [143] utilized semi-empirical calculations that employed modified Urey-Bradley and valence force fields, while Nandini and Sathyanarayana [144] employed *ab initio* Hartree-Fock calculations. The band assignments significantly differ between these studies, particularly in regard to the amide vibrations. In the work here, we performed a new normal mode analysis for propanmide using more accurate DFT calculations, and use these results to assign our infrared, normal Raman, and UVRR spectral bands.

3.4.1.1 Propanamide Molecular Structure We determined the crystal structure of our propanamide crystals (see APPENDIX B for details), and found it to be close to the structure reported previously [145]. The crystal unit cell is monoclinic ($P2_1/c$ space group), with four molecules per unit cell (Figure B1). The measured crystallographic axes and angles are a = 8.851(4), b = 5.750(2), c = 9.766(3), and β = 114.780(15)°.

We optimized the propanamide geometry (Figure 3.1) for the DFT calculations by taking the crystal structure as an initial starting point and determining the minimum energy geometry. As in the crystal, the calculated ground state equilibrium geometry of propanamide shows C_1 symmetry since the NCCC dihedral angle deviates from 180°. Our results agree with the calculated structure of Nandini and Sathyanarayana [144]. However, our calculated structure differs from that of Kuroda *et al.* [143], who assumed a C_s symmetry for

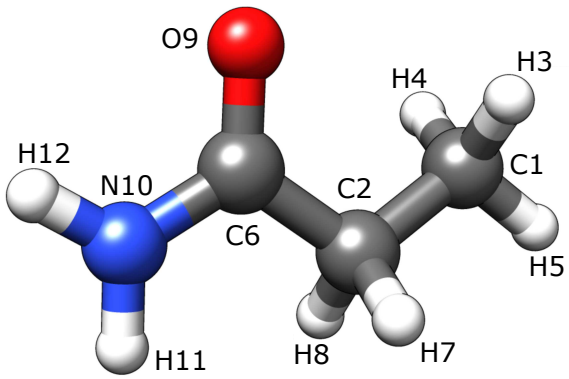


Figure 3.1: DFT-optimized structure of propanamide showing atomic numbering scheme used for normal mode analysis. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

propanamide, since its crystal structure had not been determined at that time. Propanamide has 30 fundamental vibrations, which, for a C_1 point group are all both infrared and Raman active, and unlike in the C_s point group, there is no differentiation between in-plane (a') and out-of-plane (a'') modes.

Table B4–Table B6 enable comparisons between the structural parameters of the energy minimized structure used in our calculations, and our measured crystal structure. The geometry and heavy atom bond lengths and angles of the minimized structure are close to the crystal structure, but not identical. This is most likely because the polarizable continuum model (PCM) employed in our calculations does not take into account crystal packing forces and hydrogen bonding.

Vibrational Band Assignments

Figure 3.2 shows the infrared, visible Raman, and UVRR spectra of $CH_3CH_2CONH_2$ and $CH_3CH_2COND_2$. Table 3.1 compares the frequencies and relative intensities of the infrared and non-resonance Raman spectra. We utilized UVRR to help identify the primary amide Raman bands. The primary amide NV_1 electronic transition absorption band at ~180 nm resonance enhances the 204 nm excited UVRR spectra and preresonance enhances the 229 nm excited UVRR spectra. We expect the most resonance enhanced bands will consist of vibrations with large contributions of C−N stretching, since the excited state

Table 3.1: Frequencies (cm^{-1}) of Crystalline $CH_3CH_2CONH_2$ and $CH_3CH_2COND_2$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

$CH_3CH_2CONH_2$[a]		$CH_3CH_2COND_2$[a,b]	
Infrared	Raman	Infrared	Raman
3356 vs, br	3356 s, br		
3309 sh			
	3265 sh		
3177 vs, br	3171 s, br		
2979 s	2975 s	2981 m	2977 s
2943 m	2940 vs	2943 w	2941 vs
2922 w	2909 s	2923 sh	2912 s
2882 w	2882 sh	2884 vw	2886 sh
2811 m	2827 vw		2833 sh
2737 sh	2734 m		2734 w
		2527 s	2523 s
		2407 s, br	2393 s, br
1643 sh, vs	1676 vw	1623 vs	1610 s
1628 vs	1588 m		
1464 m	1464 s	1466 m	1464 m
	1450 s		1450 sh
			1434 s
1418 vs	1420 s	1425 s	1422 sh
1379 w	1381 w	1379 w	1378 sh
1294 s	1302 w	1318 w	1317 vw
	1260 s		1261 m
		1168 vw	1176 w
1141 s	1148 s		
1068 m	1070 m	1079 m	1079 s
1007 w	1009 m	1007 w	1006 sh
		989 sh	992 m
		940 m	942 s
822 m	822 s		819 sh
811 sh	812 sh	807 m	806 m
			770 s
			731 sh
648 s,br	632 sh, br		
	563 w, br		549 w, br
	471 s		441 m
	287 vw		274 w, br
	210 vw		

[a]vs: very strong; s: strong; m: medium; w: weak; vw: very weak; br: broad; sh: shoulder. [b]Bands that derive from the mono-deuterated amide group are not reported.

Table 3.2: Frequencies (cm^{-1}) and Assignments of Infrared and Raman Bands for Crystalline $CH_3CH_2CONH_2$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

	Infrared	Raman	Calc.	δ^a (%)	PED[b] (≥5% contribution)
ν_1	3356	3356	3434	2.3	νNH (57), $-\nu$NH (43)
ν_2	3177	3171	3319	4.6	νNH (57), νNH (43)
ν_3	2979	2975	2906	2.4	νC1H (69), $-\nu$C1H (22)
ν_4	2979	2975	2894	2.8	νC1H (49), $-\nu$C1H (45)
$2\nu_{10}$	2943	2940	–		–
ν_5	2922	2909	2866	1.7	νC2H (75), $-\nu$C2H (19)
ν_6	2882	2882	2830	1.8	νC1H (42), νC1H (32), νC2H (23)
ν_7	2811	2827	2825	0.2	νC2H (76), νC2H (20)
$2\nu_{14}$	2737	2734	–		–
ν_8	1643	1676	1667	0.5	νC=O (75), $-\nu$CN (7), δ_sNC(O)C (7)
ν_9	1628	1588	1562	2.9	σNH_2 (86), νCN (10)
ν_{10}	1464	1464	1453	0.7	$\delta'_{as}CH_3$ (43), $-\delta_{as}CH_3$ (38), $\rho'CH_3$ (8)
ν_{11}	–	1450	1443	0.5	$\delta_{as}CH_3$ (51), δ_{as}'CH_3 (39), $-\rho'CH_3$ (8)
ν_{12}	–	1430[c]	1393	2.6	ωCH_2 (30), νC2C6 (20), $-\nu$CN (19), $-\beta$C=O (10), $-\nu$C1C2 (7), $\delta'_{as}CH_3$ (6)
ν_{13}	1418	1420	1422	0.2	σCH_2 (89)
ν_{14}	1379	1381	1374	0.4	$\delta_s CH_3$ (89), νC1C2 (6)
ν_{15}	1294	1302	1278	1.6	ωCH_2 (31), νCN (28), $-\tau CH_2$ (12), βC=O (8), ρNH_2 (5)
ν_{16}	–	1260	1271	0.9	τCH_2 (61), ωCH_2 (13), $-\rho'CH_3$ (7), $-\rho CH_2$ (6)
ν_{17}	1141	1148	1130	1.3	ρNH_2 (30), $-\nu$C1C2 (19), $\rho'CH_3$ (11), $-\rho CH_3$ (10), $-\nu$C=O (8), δ_sNC(O)C (7), δCCC (7)
ν_{18}	1068	1070	1108	3.6	ρCH_2 (27), $-\rho CH_3$ (21), $-\tau CH_2$ (17), $-\rho'CH_3$ (16), $-\Pi$C=O (11)
ν_{19}	1068	1070	1090	1.9	νC1C2 (38), ρNH_2 (26), $-\nu$CN (16), ρCH_3 (6)
ν_{20}	1007	1009	1028	2.0	$\rho'CH_3$ (28), ωCH_2 (19), νC1C2 (18), $-\rho CH_3$ (11), $-\nu$C2C6 (9), $-\rho CH_2$ (8)
ν_{21}	822	822	853	3.8	ρCH_2 (23), $\rho'CH_3$ (18), $-\Pi$C=O (18), ρCH_3 (15), τCH_2 (10), νC2C6 (8)
ν_{22}	811	812	850	4.7	νC2C6 (43), $-\rho CH_3$ (14), ρNH_2 (11), $-\rho CH_2$ (7), νC1C2 (6)

[a]$\delta = |\nu_{obs} - \nu_{calc}|/\nu_{obs} \times 100\%$. [b]$\nu$: stretch; δ_s: sym deformation; σ: scissoring; δ_{as}: asym deformation; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting; Π: out-of-plane bending. [c]Frequency obtained from 229 nm excitation UVRR data.

Table 3.3: Frequencies (cm^{-1}) and Assignments of Infrared and Raman Bands for Crystalline $CH_3CH_2COND_2$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

	Infrared	Raman	Calc.	δ^a (%)	PED^b (≥5% contribution)
ν_1	2981	2977	2941	1.3	νC1H (69), $-\nu$C1H (22)
ν_2	2981	2977	2929	1.7	νC1H (49), $-\nu$C1H (45)
$2\nu_{10}$	2943	2941	–	–	–
ν_3	2923	2912	2900	0.6	νC2H (75), $-\nu$C2H (19)
ν_4	2884	2886	2863	0.8	νC1H (42), νC1H (32), νC1H (23)
ν_5	–	2833	2858	0.9	νC2H (76), νC2H (20)
$2\nu_{13}$	–	2734	–	–	–
ν_6	2527	2523	2607	3.3	νND (54), $-\nu$ND (45)
ν_7	2407	2393	2458	2.4	νND (54), νND (45)
ν_8	1623	1610	1662	2.8	νC=O (78), δ_sNC(O)C (7), $-\nu$CN (7)
ν_9	1466	1464	1447	1.2	$\delta'_{as}CH_3$ (42), $-\delta_{as}CH_3$ (37), $\rho'CH_3$ (8)
ν_{10}	–	1450	1437	0.9	$\delta_{as}CH_3$ (52), $\delta'_{as}CH_3$ (38), -ρCH_3 (8)
ν_{11}	–	1434^c	1401	2.3	νCN (33), $-\omega CH_2$ (19), $-\nu$C2C6 (19), βC=O (10), $-\delta'_{as}CH_3$ (6), $\delta_s ND_2$ (5)
ν_{12}	1425	1422	1415	0.6	σCH_2 (89)
ν_{13}	1379	1378	1367	0.8	$\delta_s CH_3$ (89), νC1C2 (8)
ν_{14}	1318	1317	1295	1.7	ωCH_2 (45), νCN (23), $\delta_s ND_2$ (15),
ν_{15}	–	1261	1261	0.0	τCH_2 (75), -ρCH_3 (9), -$\rho' CH_3$ (5)
ν_{16}	1168	1176	1134	3.2	$\delta_s ND_2$ (60), βC=O (12), νC2C6 (8), $-\omega CH_2$ (5)
ν_{17}	1079	1079	1093	1.3	νC1C2 (35), ρCH_3 (19), $-\rho' CH_3$ (11), $-\delta$CCC (7), $-\delta_s$NC(O)C (5)
ν_{18}	1079	1079	1091	1.1	ρCH_2 (25), $-\rho' CH_3$ (19), ρCH_3 (17), $-\tau CH_2$ (16), $-\Pi$C=O (11), νC1C2 (7)
ν_{19}	1007	1006	1019	1.2	νC1C2 (33), $\rho' CH_3$ (20), ωCH_2 (14), $-\nu$C2C6 str (9), $-\rho CH_3$ (6), $\delta_s ND_2$ (5)
–	989	992	–	–	
ν_{20}	940	942	936	0.5	ρND_2 (34), $-\nu$CN (15), $-\rho' CH_3$ (11), δ_sNC(O)C (9), $\delta_s ND_2$ (8) ρCH_3 (6)
–	–	819	–	–	–
ν_{21}	807	806	828	2.7	ρCH_2 (29), ρCH_3 (24), $-\Pi$C=O (19), $\rho' CH_3$ (13), τCH_2 (12)
ν_{22}	–	770	776	0.7	νC2C6 (51), ρND_2 (25), $-\delta$CCC (7)

[a] $\delta = |\nu_{obs} - \nu_{calc}|/\nu_{obs} \times 100\%$. [b] ν: stretch; δ_s: sym deformation; σ: scissoring; δ_{as}: asym deformation; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting; Π: out-of-plane bending. [c] Frequency obtained from 229 nm excitation UVRR data.

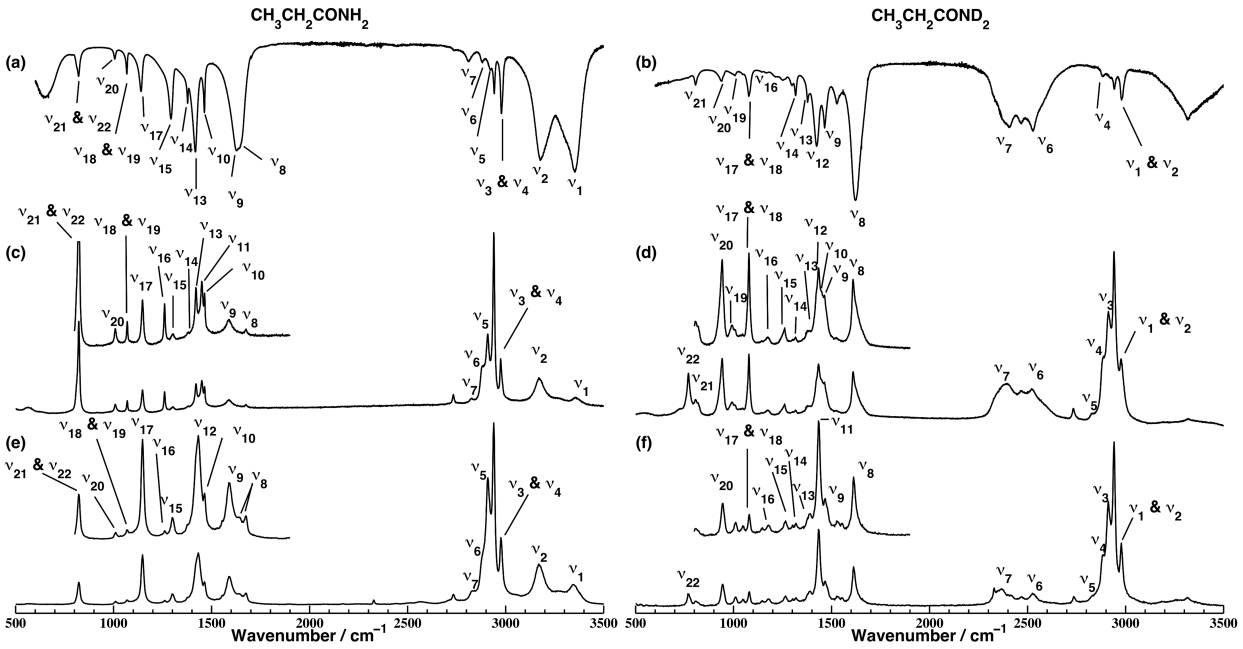


Figure 3.2: (a, b) Infrared, (c, d) non-resonance Raman, and (e, f) UVRR spectra of crystalline $CH_3CH_2CONH_2$ and $CH_3CH_2COND_2$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

is expected to be expanded along this coordinate [147]. Table 3.2 and Table 3.3 list the vibrational assignments, calculated, scaled frequencies, and PEDs for $CH_3CH_2CONH_2$ and $CH_3CH_2COND_2$, respectively.

3.4.1.2 2000-3500 cm^{-1} Region The high frequency region is dominated by N−H and C−H stretching bands. The N−H stretches show broad, strong peaks in both the infrared and Raman spectra that substantially downshift in frequency upon N-deuteration. The NH_2 asymmetric stretching band is located at ∼3355 cm^{-1}, but downshifts to ∼2525 cm^{-1} upon N-deuteration. The NH_2 symmetric stretching band appears at ∼3175 cm^{-1}, while its ND_2 counterpart appears at ∼2400 cm^{-1}. The ∼3320 cm^{-1} feature observed in the $CH_3CH_2COND_2$ spectra derives from an N−H stretching vibration for mono-N-deuterated propanamide.

The bands located between 2700–3100 cm^{-1} are insensitive to N-deuteration, which indicates that they are primarily C−H stretching modes. The ∼2975 cm^{-1} and ∼2880 cm^{-1}

bands derive from CH_3 asymmetric and symmetric stretching modes, respectively. The CH_2 asymmetric stretching mode appears at $\sim$2910 cm^{-1} in the Raman, and at $\sim$2920 cm^{-1} in the infrared spectra. In contrast, the CH_2 symmetric stretching band appears at $\sim$2830 cm^{-1} and $\sim$2810 cm^{-1} in the Raman and infrared spectra, respectively.

The bands at $\sim$2735 cm^{-1} and $\sim$2940 cm^{-1} cannot be assigned to fundamentals. Neither band shifts upon N-deuteration, and according to Kuroda *et al.*'s [143] data, these bands also do not shift upon deuteration of the methlyene group. This indicates that they do not derive from overtone or combination of amide or methlyene bands. Therefore, we assign the $\sim$2735 cm^{-1} band to the first overtone of the the $\sim$1380 cm^{-1} CH_3 symmetric deformation vibration. The $\sim$2940 cm^{-1} band is strong in both the non-resonance Raman and UVRR spectra, but is of only moderate intensity in the infrared. We assign this band to a Fermi resonance between the first overtone of the $\sim$1460 cm^{-1} CH_3 asymmetric deformation and the CH_3 symmetric stretching fundamental based on the suggestions of Kuroda *et al.* [143] and Nolin and Jones [148]. This assignment disagrees with Nandini and Sathyanarayana [144], who attribute this band to a CH_2 asymmetric stretching mode. We disagree with the Nandini and Sathyanarayana [144] assignment since this band is upshifted beyond the typical frequency range ($\sim$2910–2930 cm^{-1}) observed for CH_2 asymmetric stretching vibrations [149, 150].

3.4.1.3 1500-1800 cm^{-1} Region In this region there are two vibrations that involve the primary amide group, the Amide I (Am I) and Amide II (Am II) vibrations. The Am I at $\sim$1640 cm^{-1} and the Am II at $\sim$1620 cm^{-1} are strong and overlap in the infrared spectrum. In contrast, in the non-resonance Raman spectrum they are well-resolved, and show up as a moderately weak band located at $\sim$1675 cm^{-1} (Am I), and a stronger band at $\sim$1590 cm^{-1} (Am II). In the UVRR spectrum, the Am I band shows a $\sim$1675 cm^{-1} peak followed by an overlapping $\sim$1640 cm^{-1} feature. The Am II band is strong and occurs at $\sim$1590 cm^{-1}.

The large frequency differences between the infrared and Raman bands for the Am I and Am II vibrations presumably derive from the coupling of molecular vibrations within the crystal lattice into phonons. The spectral frequency differences for the infrared, normal Raman, and resonance Raman spectra derive from their differing selection rules for the

different phonon modes with different phasings of relative molecular motion.

Kuroda *et al.* [143] indicated that C=O stretching and NH_2 scissoring motions are both important to the PEDs of the Am I and Am II vibrations. In contrast, Nandini and Sathyanarayana's [144] normal mode analysis, as well as other studies on acetamide [151–153], indicate that the Am I vibration is mainly C=O stretching and the Am II is mainly NH_2 bending. Our analysis concludes that the Am I mode consists mostly of C=O stretching (∼75%), with minor C−N stretching and NC(O)C in-plane bending (∼7% each) components, while the Am II mode is essentially pure NH_2 scissoring (∼86%) with a small C−N stretching component (∼10%).

The spectral changes that are observed upon N-deuteration are consistent with our normal mode analysis of the Am I and Am II bands. The Am II band completely disappears, and a new band, which derives from ND_2 scissoring, appears at ∼1170 cm^{-1}, supporting the notion that this mode is essentially pure NH_2 scissoring. In contrast, N-deuteration results in the Am I mode downshifting to ∼1610 cm^{-1} in the non-resonance Raman and UVRR spectra, and to ∼1620 cm^{-1} in the infrared spectrum. This behavior is similar to the Am I band in secondary amides, and strongly supports the idea that this mode is predominately C=O stretching.

3.4.1.4 1200-1500 cm^{-1} Region Most bands in this region are easily assigned to CH_3 or CH_2 deformations and bending vibrations. The bands at ∼1464 cm^{-1} and ∼1450 cm^{-1} derive from CH_3 asymmetric deformations, while the ∼1380 cm^{-1} band is assigned to the CH_3 symmetric deformation. In the non-resonance Raman and UVRR spectra, there is a weak band at ∼1260 cm^{-1} that is assigned to a CH_2 twisting mode.

3.4.1.5 C–N Stretching Modes Our assignments of the remaining bands observed in the 1200–1500 cm^{-1} region differ from Kuroda *et al.* [143] and Nandini and Sathyanarayana [144]. Kuroda *et al.* [143] previously assigned the ∼1420 cm^{-1} band to two fundamentals, a CH_2 scissoring mode and a C−N stretching mode. Nandini and Sathyanarayana [144] also assigned this band to two fundamentals, *viz.*, a CH_2 bending mode and a CH_3 symmetric bending vibration. Nandini and Sathyanarayana [144] assigned a ∼1300 cm^{-1} band to a

vibration that consists of ∼30% C−N stretching and CH_2 wagging. They conflate this vibration with the C−N stretching mode that Kuroda *et al.* [143] assigned to the ∼1420 cm^{-1} band. In contrast, Kuroda *et al.* [143] assigns the ∼1300 cm^{-1} band to an almost pure CH_2 wagging vibration.

Our normal mode analysis and UVRR data lead to very different assignments of these two vibrations. In the UVRR spectrum of $CH_3CH_2CONH_2$, there is a very intense band at ∼1430 cm^{-1}, which we assign to the ν_{12} fundamental (Table 3.2) since it contains significant C−N stretching character. This band is not apparent in Kuroda *et al.*'s [143] or our non-resonance Raman spectra of $CH_3CH_2CONH_2$. The assignment of this resonance enhanced ∼1430 cm^{-1} mode allows us to unambiguously assign the ∼1420 cm^{-1} band to a CH_2 scissoring vibration. The ∼1300 cm^{-1} band, which appears weak in both the non-resonance Raman and UVRR spectra, is assigned to the ν_{15} vibration, which appears to be related to the ν_{12} mode.

Our normal mode analysis shows that the ν_{12} and ν_{15} amide vibrations contain significant C−N stretching. The largest ν_{12} PED components are CH_2 wagging (∼30%), C−CH_2 stretching (∼20%), C−N stretching (∼19%), and C=O in-plane bending (∼10%). For ν_{15}, the major PED components are CH_2 wagging and C−N stretching (∼30% each), followed by CH_2 twisting (∼12%) and C=O in-plane bending (∼8%).

3.4.1.6 1000-1200 cm^{-1} Region In this spectral region, we expect NH_2 rocking, CH_2 rocking, CH_3 rocking, and C−CH_3 stretching vibrations. The ∼1141 cm^{-1} infrared and ∼1148 cm^{-1} Raman bands are easily assigned to the NH_2 rocking vibration, since they downshift to ∼940 cm^{-1} upon N-deuteration. The C−CH_3 (ν_{19}) stretching vibration was previously assigned [143], but our normal mode calculation suggests that this mode is more complicated and contains significant contributions of NH_2 rocking (∼26%) and C−N stretching (∼16%). The infrared and Raman spectra show only two bands located at ∼1070 cm^{-1} and ∼1010 cm^{-1}. We assign the ∼1070 cm^{-1} band to both the CH_2 rocking (ν_{18}) and C−CH_3 stretching (ν_{19}) modes. In contrast, the ∼1010 cm^{-1} band shows a negligible change in frequency upon N-deuteration, and is thus assigned to the CH_3 rocking mode.

3.4.1.7 <1000 cm^{-1} Region The region below 1000 cm^{-1} is dominated mainly by torsional motions, as well as extensively coupled skeletal stretching and deformation modes. This region is difficult to unambiguously assign, especially below 700 cm^{-1}. We assign the two bands at ~820 cm^{-1} and 810 cm^{-1} to CH_2 rocking and C−CH_3 stretching fundamentals, respectively. The C−CH_3 stretching mode shows a modest contribution of NH_2 rocking, which likely accounts for its ~42 cm^{-1} downshift to ~770 cm^{-1} upon N-deuteration.

3.4.2 Solvation Dependence of UVRR Bands

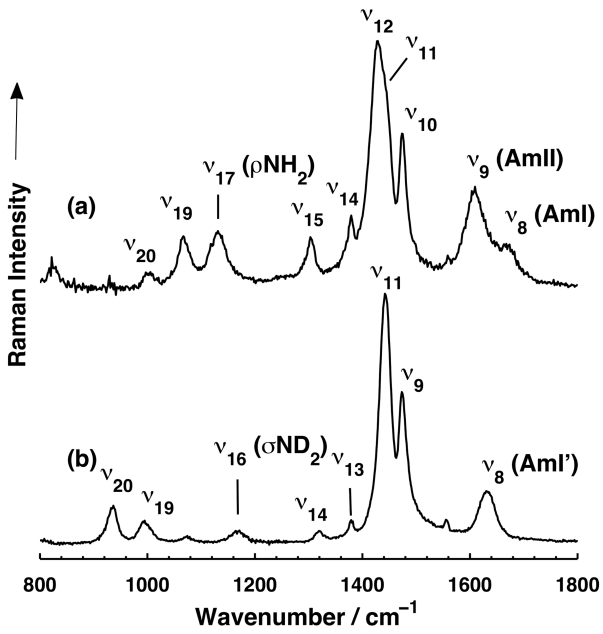


Figure 3.3: UVRR (204 nm excitation) spectra of propanamide (a) in H_2O and (b) in D_2O. The contribution of solvent was subtracted from both spectra. For (a) 200 mM $NaClO_4^-$ was used as an internal standard. The contribution of the ClO_4^- stretching band was also subtracted. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

3.4.2.1 UVRR of Propanamide in Aqueous Solutions We measured the ~204 nm excited UVRR spectra of propanamide in mixtures of acetonitrile and water in order to determine the effects of solvation on the primary amide vibrations. The spectra of propanamide in aqueous solutions (Figure 3.3) differ from that of crystalline propanamide. Compared to

Table 3.4: Measured Frequencies (cm^{-1}) and Cross Sections (σ, mbarns·molec.$^{-1}$·sr^{-1}) of UVRR bands in CH3CH2CONH$_2$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

	crystal[a]	water[b]		acetonitrile[c]	
vibration	Freq.	Freq.	σ	Freq.	σ
ν_8 (Am I)	1676	1669±1	3.5±0.70	1692	9.8
ν_9 (Am II)	1588	1610±1	18±1.1	1619	4.9
ν_{12}	1430	1428±1	30±0.90	1403	9.1
ν_{17}	1148	1132±1	8.8±0.24	1121	3.0
ν_{19}	1070	1069±1	6.0±0.14	1051	2.3

[a] UVRR cross sections were not calculated due to the lack of an internal standard. [b]Values measured from n = 4 sample size. [c]Values measured from n = 1 sample size.

crystalline spectra, propanamide in H_2O (Table 3.4) shows an Am I (ν_8) band that downshifts ~7 cm^{-1} and appears as a shoulder, while the Am II (ν_9) band upshifts ~22 cm^{-1}. The ν_{12} and C−CH_3 stretching (ν_{19}) bands do not change frequency, while the NH_2 rocking (ν_{17}) mode downshifts ~16 cm^{-1}.

The bands of propanamide in D_2O (Figure 3.3b) also show significant changes compared to their N-deuterated crystal spectra. The Am I'(ν_8) band is broad and is located at ~1633 cm^{-1}. The Am II band disappears, and the ND_2 (ν_{16}) scissoring band occurs at ~1168 cm^{-1}, while the ν_{11} band is at ~1443 cm^{-1}.

3.4.2.2 Effect of Solvation on UVRR Spectra

Figure 3.4 shows the dramatic effect of solvation on the UVRR bands of primary amides. The spectra show that all bands, except for the Am I, increase their Raman cross sections as the mole fraction of H_2O increases (Table 3.4). The Am II and ν_{12} bands show the largest cross section increases (~3–4-fold), while the NH_2 rocking (ν_{17}) and the C−CH_3 stretching (ν_{19}) bands increase ~2–3-fold. The Am I band is the only band whose Raman cross section decreases as the mole fraction of H_2O increases.

Figure 3.5 shows that there is a roughly linear cross section and frequency increase for most bands. The ν_{12} band shows a ~29 cm^{-1} per mole fraction H_2O frequency increase, the NH_2 rocking (ν_{17}) band shows a ~14 cm^{-1} per mole fraction H_2O frequency increase, while

the C−CH_3 stretching (ν_{19}) vibration shows a ∼19 cm^{-1} per mole fraction H_2O frequency increase. In contrast, the Am I band shows a ∼22 cm^{-1} per mole fraction H_2O frequency decrease. The Am II band shows only a modest frequency decrease with increasing water mole fractions.

The dependence of the UVRR spectra of propanamide on the mole fraction of water shown in Figure 3.4 is very similar to that observed in valeramide [125] and NMA [98, 154–157]. To understand this behavior, we compared the UV absorption spectra of propanamide and NMA in acetonitrile and in H_2O (Figure 3.6). The molar absorptivities of the ∼200 nm NV_1 transitions increase for both propanamide and NMA as the solvent transfers from acetonitrile to H_2O. For NMA, the absorption peak maximum of the NV_1 transition redshifts going from acetonitrile to water. This is less clearly evident in the case of propanamide where the NV_1 absorption maximum lies deeper in the UV at ∼180 nm. This trend is expected from Nielsen and Schellman's [158] results. They observe redshifts in the absorption maxima of several primary and secondary amides going from cyclohexane to water. This increase in molar absorptivity of the NV_1 transition, upon aqueous solvation, is in part responsible for increasing the UVRR cross sections due to the fact that the Raman scattering cross section is proportional to the square of the molar absorptivity.

However, most of the cross section increase results from changes in the ground state structure. The effect of aqueous solvation on the UVRR secondary amide band intensities and frequencies has been traditionally rationalized by considering the effects of the solvent dielectric and direct hydrogen bonding on the amide group resonance structures [98, 154–157]. We can invoke a similar argument for primary amides.

In low dielectric constant and hydrogen bonding environments, the O=C−NH_2 resonance form is typically dominant over the $^{-}$O−C=NH_2^+ structure in the propanamide electronic ground state (Scheme 3.1). The dipolar resonance structure becomes more favorable in water due to the high dielectric constant and the stabilizing hydrogen bonding to propanamide's C=O and NH_2 groups. These two effects increase the C−N bond order and decrease the C−O bond order of the primary amide group in the electronic ground state.

It is also important to note that the electronic excited state can also be impacted by solvation effects. For example, in NMA, Hudson and Markham [156] argued that their

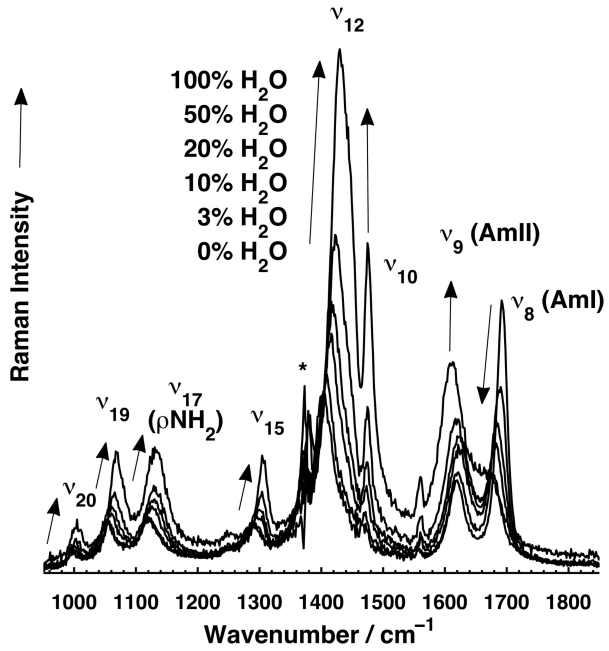


Figure 3.4: UVRR (204 nm excitation) spectra of propanamide (30 mM) in different acetonitrile and water mixtures. Arrows show frequency and intensity trends of UVRR bands as the fraction of H_2O increases. $NaClO_4^-$ (0.2 M) was used as an internal intensity and calibration standard in the solutions. Spectra were normalized to the integrated area of the 932 cm^{-1} ClO_4^- stretching band. The spectral contributions of acetonitrile, ClO_4^-, and water were subtracted. The asterisk indicates a spectral feature that is an artifact of subtracting out the ∼1376 cm^{-1} acetonitrile band. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

ab initio post-Hartree Fock calculations indicate that the effects of hydrogen bonding due to solvation are greater in magnitude on the $\pi \rightarrow \pi^*$ electronic excited state equilibrium geometry than for the ground state geometry. They argue that the changes in the C=O and C−N bond lengths in the excited state due to solvation were also in the opposite direction of the excited state bond length changes of unsolvated NMA; i.e., the C−N bond length is larger and the C=O bond length is smaller in the excited state for a $NMA(H_2O)_3$ cluster compared to an isolated NMA molecule.

The changes in C−O and C−N bond orders (and bond lengths) of the amide group

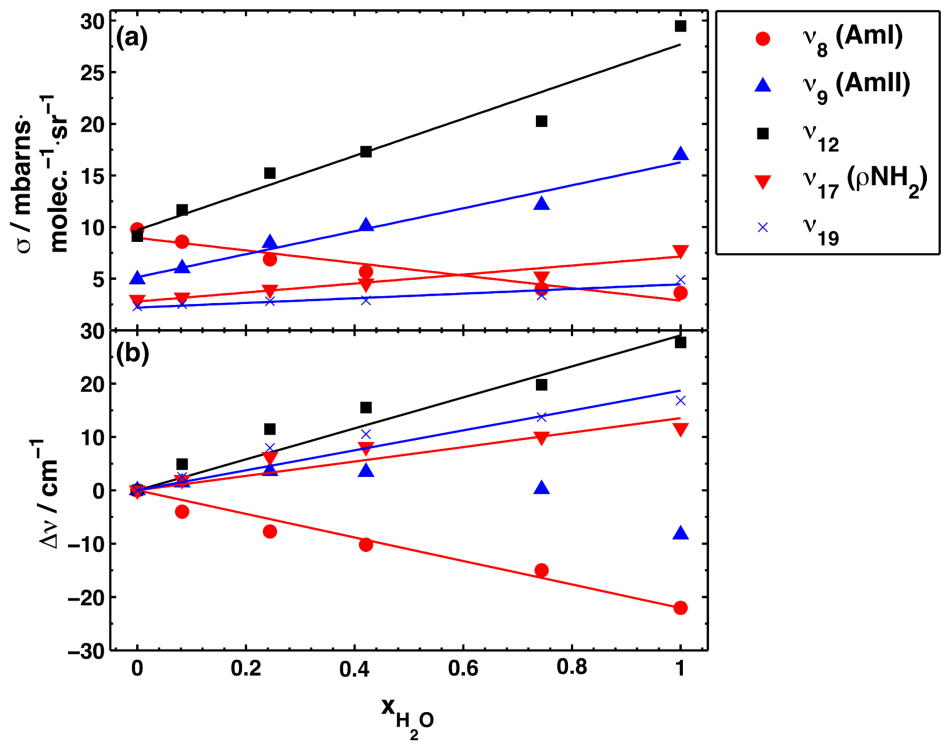


Figure 3.5: Dependence of UVRR propanamide band cross sections and frequencies on the mole fraction of water. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

going from acetonitrile to water profoundly effects the resonance Raman cross sections and the band frequencies. This is because resonance Raman cross sections scale with the square of the displacement along the enhanced vibrational normal coordinate between the equilibrium geometries of the electronic ground and excited states [159]. For example, the most resonance enhanced UVRR bands of NMA in water involve C−N stretching [98, 154, 155, 160, 161] because of the large expansion of the electronic excited state along the C−N bond [156, 157, 161]. In contrast, there is a relatively small enhancement of the Am I band for NMA in water because the excited state expansion along the C=O bond is much less.

In the case of propanamide, the dramatic spectral changes observed in Figure 3.4 can likewise be explained by changes in C=O and C−N bond lengths upon solvation. The elongation of the C=O bond results in a decrease of the C=O stretching force constant. This results in a downshift in the Am I band frequency. The Am I band UVRR cross section

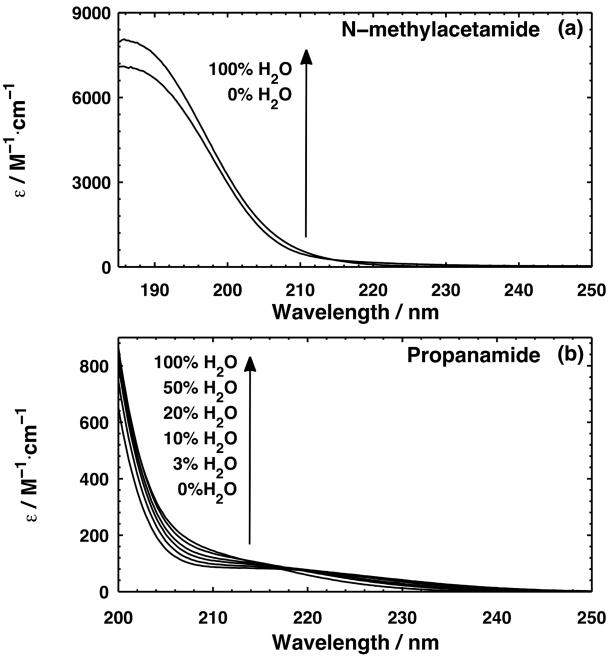
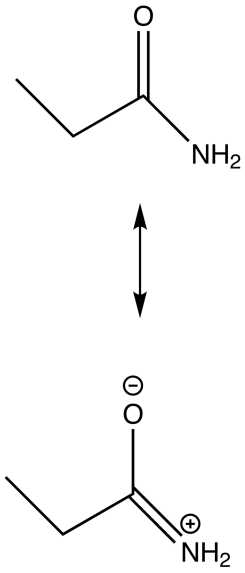


Figure 3.6: UV absorption spectra of (a) NMA and (b) propanamide in acetonitrile and water. Arrows show trend of molar absorptivity as the fraction of water increases. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.



Scheme 3.1: Resonance Structures of Propanamide. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

also decreases because the magnitude of the displacement between the electronic ground and excited states along the C−O coordinate must decrease in water since the ground state C=O bond length elongates.

In contrast, for vibrations with significant C−N stretching, the contraction of the C−N bond length results in a vibrational frequency upshift due to the increase in the stretching force constant. This bond contraction also increases the magnitude of the displacement between the electronic ground and excited states along the C−N coordinate, which results in an increase in the Raman cross sections. It will be more difficult to explain the origin of the UVRR cross section increases for heavily coupled modes, such as the $\sim$1430 cm^{-1} ν_{12} and the $\sim$1300 cm^{-1} ν_{15} modes, which contain displacements of multiple atoms. The displacements of all of the primary amide atoms contribute to resonance enhancement. Phasing of this motion can be very important [161]. However, we conclude that the C−N stretching motion in these vibrations is predominantly responsible for their UVRR intensity enhancements.

3.5 CONCLUSION

We utilized DFT calculations, infrared, non-resonance Raman, and UVRR spectra to assign the vibrational bands of crystalline propanamide. Our study resolves previous inconsistencies in the vibrational assignments and normal mode compositions of primary amide bands. We also studied the effect of aqueous solvation on the primary amide UVRR bands by examining the $\sim$204 nm UVRR dependence as the solvent transfers from acetonitrile to water. The aqueous solvation dependence of primary amide UVRR bands can be rationalized by the stabilization between dipolar resonance structures of the ground electronic state of the amide group. Both hydrogen bonding interactions and the increased dielectric constant as the solvent transfers from acetonitrile to water contribute to the stabilization of the dipolar resonance structure in the ground state, which effectively increases the C−N bond order while decreasing the C−O bond order.

The resulting increased displacement between electronic ground and excited state geometries along the C−N coordinate increases the UVRR cross sections of vibrations that contain

significant C−N stretching. In contrast, the decreased displacement between the electronic ground and excited state geometries along the C−O coordinate results in a dramatic decrease in the Am I band UVRR cross section. These results indicate that the Am I, Am II, and ν_{12} band UVRR cross sections and their frequencies can be used as sensitive spectroscopic markers for hydrogen bonding and local dielectric environment of the side chains of Gln and Asn.

3.6 ACKNOWLEDGMENT

Funding for this work was provided by the University of Pittsburgh. E.M.D. gratefully acknowledges support through the NIH Molecular Biophysics and Structural Biology Training Grant (T32 GM88119-3). The computational work was supported by the University of Pittsburgh Center for Simulation and Modeling through the super-computing resources provided.

4.0 GLUTAMINE AND ASPARAGINE SIDE CHAIN HYPERCONJUGATION INDUCED STRUCTURALLY SENSITIVE VIBRATIONS

Adapted with permission from: David Punihaole, Zhenmin Hong, Ryan S. Jakubek, Elizabeth M. Dahlburg, Steven Geib, Sanford A. Asher. Glutamine and Asparagine Side Chain Hyperconjugation Induced Structurally Sensitive Vibrations. *The Journal of Physical Chemistry B*, **2015**, 119, 13039–13051.

Author Contributions: D.P. acquired and analyzed the UVRR data with the assistance of R.S.J. and E.M.D. D.P. and S.G. acquired and analyzed X-ray diffraction data. Z.H. performed and analyzed the results of the DFT calculations with the assistance of D.P. The manuscript was prepared by D.P. and S.A.A. with the assistance of R.S.J. and E.M.D.

We identified vibrational spectral marker bands that sensitively report on the side chain structures of Gln and Asn. DFT calculations indicate that the Amide III^P (Am III^P) vibrations of Gln and Asn depend cosinusoidally on their side chain OCCC dihedral angles (the χ_3 and χ_2 angles of Gln and Asn, respectively). We use UVRR and visible Raman spectroscopy to experimentally correlate the Am III^P Raman band frequency to the primary amide OCCC dihedral angle. The Am III^P structural sensitivity derives from the Gln (Asn) C_β–C_γ (C_α–C_β) stretching component of the vibration. The C_β–C_γ (C_α–C_β) bond length inversely correlates with the Am III^P band frequency. As the C_β–C_γ (C_α–C_β) bond length decreases, its stretching force constant increases, which results in an upshift in the

Am III^P frequency. The C_β–C_γ (C_α–C_β) bond length dependence on the χ_3 (χ_2) dihedral angle results from hyperconjugation between the C_δ=O_ϵ (C_γ=O_δ) π^* and C_β–C_γ (C_α–C_β) σ orbitals. Using a protein data bank library, we show that the χ_3 and χ_2 dihedral angles of Gln and Asn depend on the peptide backbone Ramachandran angles. We demonstrate that the inhomogeneously broadened Am III^P band line shapes can be used to calculate the χ_3 and χ_2 angle distributions of peptides. The spectral correlations determined in this study enable important new insights into protein structure in solution, and in Gln- and Asn-rich amyloid-like fibrils and prions.

4.1 INTRODUCTION

Amyloid-like fibril protein aggregates and prion proteins often contain stretches of Gln and Asn residues. For example, polyQ-rich fibrils are the pathological hallmarks of several "CAG" codon repeat diseases [1, 3, 4, 35, 45, 162–165]. Similarly, Sup35p and Ure2p prions contain Gln- and Asn-rich regions that drive their aggregation and cause loss-of-function of these normally soluble proteins [166].

Because Gln and Asn side chains can hydrogen bond to water, the peptide backbone, or other side chains, they serve unique roles in protein structure and conformational transitions. Unfortunately, there is relatively little known about the mechanisms by which the primary amide groups of Gln and Asn interact with other protein constituents, or what role they play in the aggregation of prions and fibrils. Consequently, it is important to find spectroscopic markers that can be used to monitor the conformations and hydrogen bonding environments of Asn and Gln side chains in order to develop a deeper understanding of the roles that these residues play in protein aggregation.

There are few methods to quantitatively examine the conformations of Gln and Asn side chains in prion and fibril aggregates. Recent solid-state NMR studies [56, 58, 59] suggest that there are at least two different populations of Gln side chain conformers in polyQ fibrils. Sharma et al. [55] claim on the basis of low-resolution X-ray fiber and powder diffraction data that the side chains in polyQ fibrils adopt an unusual bent conformation; however,

these highly uncommon side chain structures have not been substantiated by other studies.

High resolution X-ray diffraction studies [167, 168] on small peptide microcrystals that contain amyloidogenic sequences have revealed important, atomic-resolution details regarding the steric zipper interactions that could occur in Gln- and Asn-rich prions and fibrils. These studies indicate, for example, that differences in the structures and hydrogen bonding interactions of amino acid side chains give rise to different fibril polymorphs. However, the conformations observed in the small peptide crystals may not reflect the side chain structures and hydrogen bonding interactions that occur in *bona fide* prion and fibril aggregates.

UVRR is a powerful, emerging tool for studying the conformations of proteins, as well as the structure, local hydrogen bonding, and dielectric environments of amino acid side chains [89, 101–103, 123, 124, 146, 169–173]. Deep UV excitation (∼200 nm) selectively resonance enhances secondary and primary amide vibrations [125, 134]. Previous investigations of secondary amide vibrations have developed a detailed understanding of the UVRR spectral dependence on the peptide bond structure and its hydrogen bonding [98, 155, 156, 174]. For example, Asher and coworkers [96, 97, 105] quantitatively correlated the Amide III_3 (Am III_3) frequency to the peptide bond Ramachandran Ψ dihedral angle. They determined that the structural sensitivity of the Am III_3 vibration derives from coupling between the peptide backbone amide N–H and C_α–H bending motions [96]. This fundamental insight enabled incisive investigations that elucidated, in detail, the mechanism of α-helix (un)folding in a wide range of solution environments [106, 107, 175–178].

We seek to develop a similar deep understanding of the UVRR spectral dependence of primary amide vibrations on the structure of Gln and Asn side chains. In this work, we discover the structural sensitivity of the Amide III^P (Am III^P) vibration on the primary amide OCCC dihedral angle (the χ_3 and χ_2 angles of the side chains Gln and Asn, respectively). The potential energy distribution (PED) of this vibration in Gln (Asn) contains significant contributions of C_β–C_γ (C_α–C_β) stretching, $N_\epsilon H_2$ ($N_\delta H_2$) rocking, and C_δ–N_ϵ (C_γ–N_δ) stretching motions. We find that the structural sensitivity of the Am III^P mode originates mainly from the C_β–C_γ (C_α–C_β) bond length dependence on the χ_3 (χ_2) dihedral angle. We demonstrate that the C_β–C_γ (C_α–C_β) bond length correlation on the χ_3 (χ_2) dihedral angle derives from hyperconjugation between the C_β–C_γ (C_α–C_β) σ orbital and the

C_δ=O_ϵ (C_γ=O_δ) π^* orbital.

We compare our results with the Gln and Asn entries of the Shapovalov and Dunbrack side chain rotamer library [179] and examine the dependence of Gln (Asn) χ_3 (χ_2) dihedral angles on the peptide backbone Ramachandran (Φ, Ψ) angles. We observe distinct χ_3 and χ_2 dihedral angle preferences for Gln and Asn residues that adopt PPII, β-sheet, and α-helix Ramachandran angles. Applying this new insight, and the dependence of the χ_3 dihedral angle on the Am III^P vibrational frequency, we determine the χ_3 angle distribution of Q_3 and $D_2Q_{10}K_2$ peptides in aqueous solution. We find that Q_3 and $D_2Q_{10}K_2$ favor χ_3 dihedral angles similar to those of Gln in solution. This result is consistent with Q_3 and $D_2Q_{10}K_2$ containing side chains that are completely solvated.

Our work here develops a novel spectral marker for experimentally probing the structures of Asn and Gln side chains in fibrils and prion aggregates. Our methodology does not require extensive isotopic labeling or crystallization and allows us to monitor the side chain structural changes that occur during protein aggregation. This enables crucial, molecular-level insights into the role that Gln and Asn side chains play in stabilizing fibril and prion aggregates. We are developing new insights into why Gln- and Asn-rich sequences have strong propensities to aggregate into amyloid-like fibrils and prions.

4.2 EXPERIMENTAL DETAILS

4.2.1 Materials

L-glutamine (L-Gln, ≥99% purity), L-glutamine t-butyl ester hydrochloride (GlnTBE, ≥98% purity), and glycyl-L-glutamine (Gly-Gln, ≥97% purity) were purchased from Sigma-Aldrich. D-glutamine (D-Gln, 98% purity) was purchased from Acros Organics, and N-acetyl-L-glutamine (NAcGln, 97% purity) was purchased from Spectrum Chemical Mfg. Corp. L-seryl-L-asparagine (Ser-Asn, ≥99% purity) was purchased from Bachem. Optima-grade H_2O was purchased from Fisher Scientific, and D_2O (99.9% atom D purity) was purchased from Cambridge Isotope Laboratories, Inc. Gln_3 was purchased from Pierce Biotech-

nology at 95% purity.

4.2.2 Sample Preparation

Gly-Gln and Ser-Asn were obtained as crystalline powders and used without further purification or re-crystallization. D-Gln, NAcGln, L-Asn, and GlnTBE crystals were prepared by drying saturated solutions in water. L-Gln crystals were obtained by drying a saturated solution in the presence of 0.1 M NaCl. N-deuterated crystals were prepared via multiple rounds of re-crystallization in D_2O. Samples of Gln_3 were prepared at $0.5\,mg\,mL^{-1}$ in HPLC-grade water containing 0.05 mM sodium perchlorate (Sigma Aldrich, ≥98% purity). The sodium perchlorate was used as an internal intensity standard to allow us to subtract the contribution of water.

4.2.3 X-ray Diffraction

X-ray diffraction of crystals was performed using a Bruker X8 Prospector Ultra equipped with a copper micro-focus tube ($\lambda = 1.541\,78\,Å$). The crystals were mounted and placed in a cold stream of N_2 gas (230 K) for data collection. The frames collected on each crystal specimen were integrated with the Bruker SAINT software package using the narrow-frame algorithm. APPENDIX C discusses, in detail, the methods used to determine the unit cells and crystal structures of the compounds examined.

4.2.4 Visible Raman Spectroscopy

Visible excitation Raman spectra of crystals were collected using a Renishaw inVia spectrometer equipped with a research-grade Leica microscope. Spectra were collected using the 633 nm excitation line from a HeNe laser and a 5× objective lens. The spectrometer resolution was $\sim 2\,cm^{-1}$. The $918\,cm^{-1}$ and $1376\,cm^{-1}$ bands of acetonitrile [129] were used to calibrate the spectral frequencies.

4.2.5 UVRR Spectroscopy

UVRR spectra of crystals were collected using CW 229 nm light generated by an Innova 300C FreD frequency doubled Ar^+ laser [130]. The crystalline specimens were spun in a cylindrical brass cell to prevent the accumulation of thermal or photodegradation products. A SPEX triplemate spectrograph, modified for use in the deep UV, was utilized to disperse the Raman scattered light. A Spec-10 CCD camera (Princeton Instruments, Model 735-0001) with a Lumagen-E coating was used to detect the Raman light. The power of UV light illuminating the sample ranged from ∼1.5–2 mW. The 801 cm^{-1}, 1028 cm^{-1}, 2852 cm^{-1}, and the 2938 cm^{-1} bands of cyclohexane were used to calibrate the 229 nm excitation UVRR spectral frequencies.

UVRR solution-state measurements were made using ∼204 nm nm excitation. The UV light was generated by Raman shifting the third harmonic of a Nd:YAG laser (Coherent, Inc.) with H_2 gas (∼30 psi) and selecting the fifth anti-Stokes line. A thermostatted (20 °C) flow cell was employed to circulate solutions in order to prevent the contribution of photodegradation products. The scattered light was dispersed and imaged using a double monochromator, modified for use in the UV in a subtractive configuration [131], and detected with a Spec-10 CCD camera.

4.3 COMPUTATIONAL DETAILS

4.3.1 Density Functional Theory (DFT) Calculations

The DFT calculations [135] were carried out using the GAUSSIAN 09 package [136]. The geometry optimizations and frequency calculations were performed using the M06-2X functional [137] and the 6-311++g** basis set. The presence of water was simulated implicitly by employing a polarizable continuum dielectric model (PCM). Vibrational frequencies were calculated using the harmonic approximation. The calculated frequencies were not scaled. The potential energy distribution (PED) of each vibration was obtained from the GAUSSIAN output files by employing a MATLAB program that we wrote. Figure 4.1 shows

the DFT calculated minimum energy structure of L-Gln and the atomic labeling scheme. In order to study the conformational dependence of the Raman bands, we fixed the χ_3 dihedral angle of L-Gln, re-optimized the geometry, and calculated the harmonic vibrational frequencies for a series of conformers with χ_3 angles of $-16°$, $0°$, $4°$, $\pm30°$, $\pm60°$, $\pm90°$, $\pm120°$, $\pm150°$, and $\pm180°$.

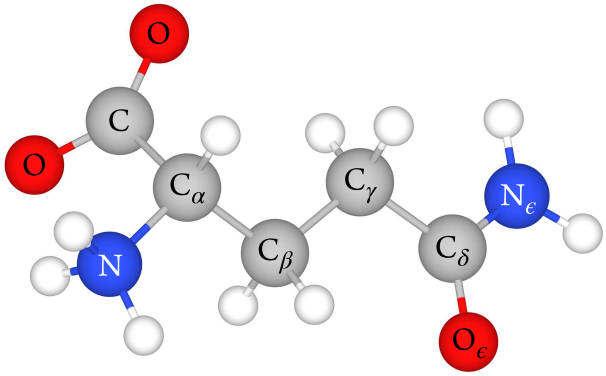


Figure 4.1: Geometry of optimized structure and atomic labeling scheme of L-Gln used in DFT calculations and band assignments. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

4.4 RESULTS AND DISCUSSION

4.4.1 Assignment of L-Gln UVRR Bands in H_2O and D_2O

Figure 4.2 shows the band-resolved ~204 nm excitation UVRR spectra of L-Gln in H_2O and D_2O. Visible Raman and infrared spectra band assignments of L-Gln were reported previously by Ramirez and coworkers for both solid-state crystalline samples [181] and in the solution-state [182]. Ramirez and coworkers made assignments without performing normal mode calculations in their study of crystalline L-Gln. In their solution-state study, they employed DFT calculations to aid in band assignments. We found that their reported frequencies did not match those in our solution-state UVRR spectra, and that their band assignments were inconsistent with our intensity expectations of the resonance enhanced

bands.

In the work here, we perform a new normal mode analysis of L-Gln in order to assign our UVRR spectra. We employ DFT calculations that use a more modern functional (M06-2X) than that of Ramirez and coworkers. These assignments build off of our previous, detailed assignment of propanamide [146], a model for the side chains of Gln and Asn. Our assignments of L-Gln in H_2O and D_2O are shown in Table 4.1 and Table 4.2, respectively.

The UVRR spectra are dominated by bands that derive from vibrations of the primary amide group. This is because these resonance enhanced vibrations couple to the strong $\sim$180 nm NV_1 electronic transition. These resonance enhanced amide bands contain significant contributions of $C_\delta{-}N_\epsilon$ stretching, since the electronic excited state is expanded along this coordinate [147].

The spectral region between 1600 cm^{-1} and 1700 cm^{-1} is dominated by two primary amide vibrations, the Amide I^P (Am I^P) and Amide II^P (Am II^P) bands. The superscript ^{P} denotes the primary amide to distinguish these vibrations from the widely known vibrations of secondary amides found in proteins. The Am I^P band is located at $\sim$1680 cm^{-1} and derives mainly from $C_\delta{=}O_\epsilon$ stretching. In D_2O, the Am I^P band (called the AmI'^P) downshifts to $\sim$1650 cm^{-1}. The Am II^P band at $\sim$1620 cm^{-1} derives from a vibration whose PED contains mostly $N_\epsilon H_2$ scissoring ($\sim$86%) and C_δ–N_ϵ stretching ($\sim$10%). Upon N-deuteration, the C_δ–N_ϵ stretching and ND_2 scissoring motions decouple. This causes the Am II^P band to disappear, and a new band, which derives from $N_\epsilon D_2$ scissoring, appears at $\sim$1160 cm^{-1}.

The most intense features of the Gln spectra in Figure 4.2 occur in the region between 1400 cm^{-1} to 1500 cm^{-1}. Most of the bands found in this region derive from CH_2 scissoring or wagging modes. However, we assign the most intense band, located at $\sim$1430 cm^{-1}, to a vibration that contains significant contributions of CH_2 wagging, C_γ–C_δ stretching, CH_2 scissoring, and C_δ–N_ϵ stretching in its PED. This assignment is based on our previous work with propanamide [146], which shows a similar intense band at $\sim$1430 cm^{-1}.

The region between 1200 cm^{-1} to 1400 cm^{-1} contains bands that derive mostly from $C_\alpha H$ rocking, CH_2 wagging, and CH_2 twisting modes. We assign the $\sim$1365 cm^{-1} and $\sim$1350 cm^{-1} bands in the Figure 4.2a spectrum to $C_\alpha H$ rocking modes. We assign the strong bands located at $\sim$1330 cm^{-1}, $\sim$1290 cm^{-1}, and the very weak $\sim$1205 cm^{-1} bands to CH_2 twisting

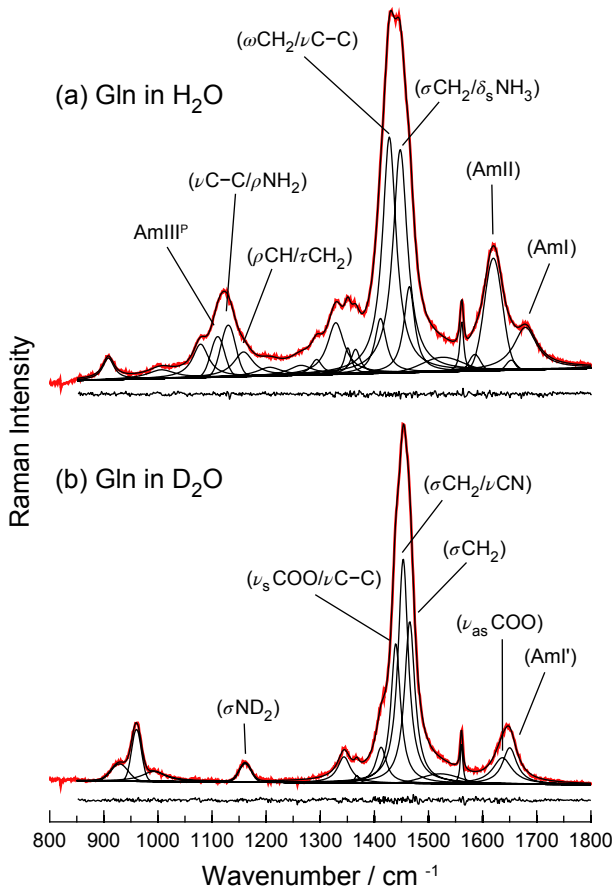


Figure 4.2: UVRR spectra excited at ∼204 nm of L-Gln in (a) H_2O and (b) D_2O. The spectral contributions of the solvents have been subtracted. The reduced χ^2 (χ^2_{red}) statistics for the spectral fits shown in (a) and (b) are 0.55 and 1.8, respectively. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

modes. The $\sim$1265 cm^{-1} feature is assigned to a CH_2 wagging vibration. Only two bands, at $\sim$1370 cm^{-1} and $\sim$1345 cm^{-1}, appear in D_2O. We assign the $\sim$1370 cm^{-1} band to a CH_2 wagging mode and the $\sim$1345 cm^{-1} band to a $C_\alpha H$ rocking mode. We conclude that these vibrations appear strongly in the UVRR spectrum in Figure 4.2b because they contain significant C_δ–N_ϵ stretching.

The region between 1000 cm^{-1} and 1200 cm^{-1} contains bands that derive from vibrations with large C–C stretching, $N_\epsilon H_2$ rocking, or NH_3 rocking contributions. Most of the vibrations in this region are complex. We assign the $\sim$1160 cm^{-1} band to a coupled $C_\alpha H$

Table 4.1: UVRR Frequencies (cm^{-1}) and Assignments of L-Gln in H_2O. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

Expt.	Calc.	Potential Energy Distribution[a] (≥5% contribution)
1679	1745	$\nu C_\delta O_\epsilon$ (75), $-\nu C_\delta N_\epsilon$ (8), $\beta N_\epsilon C_\delta C_\gamma$ (7)
1652	1666	$\delta_{as}'NH_3$ (48), $-\delta_{as}NH_3$ (44), $-\rho NH_3$ (5)
1620	1623	$-\sigma N_\epsilon H_2$ (86), $-\nu C_\delta N_\epsilon$ (10)
	1621	$\delta_{as}NH_3$ (40), $\delta_{as}'NH_3$ (48), $\delta_s NH_3$ (10)
1585	1715	$-\nu CO$ (53), νCO (33), $\rho C_\alpha C$ (7)
1464	1494	$\sigma C_\beta H_2$ (88)
	1466	$-\delta_s NH_3$ (33), $\sigma C_\gamma H_2$ (30), $\omega C_\gamma H_2$ (7), $-\omega C_\beta H_2$ (5), $\nu C_\beta C_\gamma$ (5)
1447	1460	$-\sigma C_\gamma H_2$ (44), $-\delta_s NH_3$ (36)
1427	1445	$-\omega C_\gamma H_2$ (17), $\nu C_\gamma C_\delta$ (13), $\sigma C_\gamma H_2$ (13), $\omega C_\beta H_2$ (13), $-\nu C_\delta N_\epsilon$ (9), $-\nu C_\beta C_\gamma$ (5), $-\delta_s NH_3$ (5), $\rho C_\delta O$ (5)
1411	1419	$-\nu CO$ (26), $\rho C_\alpha H$ (12), $\nu C_\alpha C$(11), $\omega C_\gamma H_2$ (8), βCOO (7), $\nu C_\delta N_\epsilon$ (7), $-\nu CO$ (7), $-\nu C_\gamma C_\delta$ (6)
1365	1388	$\rho C_\alpha H$ (25), $\omega C_\beta H_2$ (13), νCO (10), $-\rho' C_\alpha H$ (9), $\nu C_\alpha C_\beta$ (8), $-\tau C_\beta H_2$ (6), $-\nu C_\alpha C$(5)
1351	1358	$-\rho C_\alpha H$ (30), $\omega C_\beta H_2$ (16), $-\tau C_\beta H_2$ (15), $\tau C_\gamma H_2$ (9), $\nu C_\delta N_\epsilon$ (6)
1328	1348	$-\tau C_\beta H_2$ (26), $-\omega C_\beta H_2$ (18), $-\rho' C_\alpha H$ (13), $-\nu C_\delta N_\epsilon$ (11)
1293	1309	$-\tau C_\gamma H_2$ (35), $-\rho' C_\alpha H$ (26), $\rho C_\beta H_2$ (8), $-\rho C_\alpha H$ (6)
1264	1272	$-\omega C_\gamma H_2$ (43), $-\omega C_\beta H_2$ (20), $\nu C_\delta N_\epsilon$ (13)
1206	1215	$-\tau C_\beta H_2$ (21), $-\nu C_\alpha C_\beta$ (18), $-\tau C_\gamma H_2$ (16), $-\rho' NH_3$ (13), $\delta NC_\alpha C(OO)$ (5)
1158	1153	$-\rho' C_\alpha H$ (20), $\tau C_\gamma H_2$ (17), $-\rho' NH_3$ (13), $\tau C_\beta H_2$ (12), $-\nu C_\alpha C_\beta$ (8), $\rho C_\gamma H_2$ (6)
1130	1122	$\nu C_\beta C_\gamma$ (34), $\rho N_\epsilon H_2$ (17), $-\nu C_\alpha C_\beta$ (7), $\nu C_\alpha N$ (6), $-\beta N_\epsilon C_\delta C_\gamma$ (5)
	1109	ρNH_3 (27), $-\rho' C_\alpha H$ (10), $-\rho N_\epsilon H_2$ (10), $-\delta'NCC(OO)$ (9), $-\rho C_\alpha H$ (7), $\nu C_\alpha N$ (7)
1110	1097	$\nu C_\beta C_\gamma$ (26), $-\rho N_\epsilon H_2$ (26), $-\nu C_\delta N_\epsilon$ (13), $-\rho NH_3$ (8)
	1038	$\nu C_\alpha N$ (36), $-\nu C_\beta C_\gamma$ (9), $\rho C_\beta H_2$ (8), $\rho' NH_3$ (6), $\rho C_\gamma H_2$ (5), $\rho' C_\alpha H$ (5)
1078	1003	ρNH_3 (25), $-\nu C_\alpha N$ (19), $\rho C_\gamma H_2$ (14), $\rho C_\beta H_2$ (14), $\nu C_\alpha C_\beta$ (7), $-\tau C_\beta H_2$ (5)
1006	974	$-\rho' NH_3$ (38), $\nu C_\alpha C_\beta$ (25), $-\nu C_\alpha C$ (8), $-\sigma CC_\alpha C_\beta$ (7), $\nu C_\alpha N$ (6)

[a] ν: stretch; δ_{as}: asymmetric deformation; δ_s: symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

Table 4.2: UVRR Frequencies (cm^{-1}) and Assignments of L-Gln in D_2O. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

Expt.	Calc.	Potential Energy Distribution[a] (≥5% contribution)
1650	1739	$\nu C_\delta O_\epsilon$ (78), $-\nu C_\delta N_\epsilon$ (7), $\beta N_\epsilon C_\delta C_\gamma$ (7)
1637	1708	νCO (55), $-\nu$CO (34), $-\rho C_\alpha C$ (7)
1465	1494	$\sigma C_\beta H_2$ (88)
1453	1465	$-\sigma C_\gamma H_2$ (48), $-\omega C_\gamma H_2$ (15), $-\nu C_\beta C_\gamma$ (8), $-\nu C_\delta N_\epsilon$ (7), $\omega C_\beta H_2$ (6)
1440	1454	$-\sigma C_\gamma H_2$ (35), $\nu C_\delta N_\epsilon$ (19), $-\nu C_\gamma C_\delta$ (14), $\omega C_\gamma H_2$ (9), $-\rho C_\delta O_\epsilon$ (5)
1412	1424	$-\nu$CO (29), $\nu C_\alpha C$ (12), $-\nu$CO (10), $\omega C_\beta H_2$ (10), βCOO (9), $\nu C_\delta N_\epsilon$ (9), $\rho C_\alpha H$ (8)
1368	1393	$\omega C_\beta H_2$ (21), $\rho C_\alpha H$ (15), νCO (11), $\nu C_\alpha C_\beta$(8), $-\rho' C_\alpha H$ (6), νCO (6), $-\nu C_\alpha C$ (5), $\nu C_\delta N_\epsilon$ (5)
1344	1362	$-\rho C_\alpha H$ (37), $-\omega C_\gamma H_2$ (13), $\nu C_\delta N_\epsilon$ (12), $\omega C_\beta H_2$ (11), $\rho' C_\alpha H$ (7), $\sigma N_\epsilon D_2$ (5)
	1349	$-\tau C_\beta H_2$ (44), $\tau C_\gamma H_2$ (20), $-\rho C_\alpha H$ (15), $-\rho' C_\alpha H$ (6)
	1304	$-\rho' C_\alpha H$ (32), $-\tau C_\gamma H_2$ (29), $\rho C_\beta H_2$ (8), $-\rho C_\alpha H$ (7)
	1282	$-\omega C_\gamma H_2$ (37), $-\omega C_\beta H_2$ (33), $\sigma N_\epsilon D_2$ (8), $\rho C_\alpha H$ (7), $\nu C_\delta N_\epsilon$ (6)
	1202	$-\delta_{as} ND_3$ (23), $\delta_{as}' ND_3$ (20), $\tau C_\beta H_2$ (17), $\tau C_\gamma H_2$ (12), $\nu C_\alpha C_\beta$(8)
	1198	$\delta_{as} ND_3$ (27), $\tau C_\beta H_2$ (12), $\tau C_\gamma H_2$ (11), $-\rho' C_\alpha H$ (9), $\nu C_\alpha C_\beta$(8), $\delta_s ND_3$ (7), $-\delta_{as}' ND_3$ (7)
	1186	$-\delta_{as}' ND_3$ (52), $-\delta_s ND_3$ (27), $-\nu C_\alpha N$ (8), $-\delta_{as} ND_3$ (6)
1161	1150	$\sigma N_\epsilon D_2$ (56), $\rho' C_\delta O_\epsilon$ (11), $\nu C_\gamma C_\delta$ (9), $\omega C_\gamma H_2$ (8)
	1131	$\nu C_\alpha C_\beta$(21), $-\delta_s ND_3$ (16), $-\nu C_\alpha N$ (14), $\delta_{as} ND_3$ (13), $-\tau C_\gamma H_2$ (5)
	1119	$-\nu C_\beta C_\gamma$ (17), $\delta_{as} ND_3$ (10), $\delta NC_\alpha C$(OO) (9), $-\nu C_\alpha C_\beta$(8), $-\delta_s ND_3$ (8), $\sigma C_\gamma C_\beta C_\alpha$ (7), $-\nu C_\alpha N$ (6), $\delta_{as}' ND_3$ (5)
	1105	$\nu C_\beta C_\gamma$ (44), $-\delta_s ND_3$ (15), $\delta_{as} ND_3$ (8), $\delta_{as}' ND_3$ (6), $-\nu C_\alpha C_\beta$(5)
	1058	$\rho C_\beta H_2$ (21), $\rho' C_\alpha H$ (16), $\rho C_\gamma H_2$ (12), $\delta' NC_\alpha C$(OO) (9), $-\nu C_\alpha C$ (6), $\delta C' C_\alpha C_\beta$ (6), $\nu C_\alpha C_\beta$(5), $-\rho ND_3$ (5)
992	998	$\nu C_\alpha N$ (29), $\delta_s ND_3$ (15), $-\delta' NC_\alpha C$(OO) (10), $-\rho C_\alpha C$ (7), $-\nu C_\beta C_\gamma$ (5), $-\sigma CC_\alpha C_\beta$ (5)
960	959	$\rho N_\epsilon D_2$ (22), $\nu C_\delta N_\epsilon$ (14), $\nu C_\gamma C_\delta$ (10), $-\sigma N_\epsilon D_2$ (9), $-\nu C_\alpha N$ (9), $-\beta N_\epsilon C_\delta C_\gamma$ (8), $\nu C_\alpha C$ (6), $\nu C_\delta O_\epsilon$ (6)
929	920	$\nu C_\alpha C$ (17), $\rho C_\gamma H_2$ (16), ρND_3 (16), $-\tau C_\beta H_2$ (8), βCOO (8), $-\delta' NC_\alpha C$(OO) (7), $\Pi N_\epsilon C_\delta C_\gamma$ (6)

[a] ν: stretch; δ_{as}: asymmetric deformation; δ_s: symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting; Π: out-of-plane deformation.

rocking/CH_2 twisting mode. The PED of this vibration contains a significant contribution of NH_3 rocking, which likely accounts for the disappearance of this band upon N-deuteration. We assign the $\sim$1080 cm^{-1} and $\sim$1005 cm^{-1} bands to NH_3 rocking vibrations.

The remaining two bands in the 1000 cm^{-1} to 1200 cm^{-1} region are located at $\sim$1130 cm^{-1} and $\sim$1110 cm^{-1}. The observed frequency difference between these two vibrations is $\sim$20 cm^{-1}, which is close to the calculated $\sim$25 cm^{-1} difference of our DFT calculations. We assign the 1130 cm^{-1} band to a vibration that is mainly an in-phase combination of C_β–C_γ stretching and $N_\epsilon H_2$ rocking. The $\sim$1110 cm^{-1} band is assigned to a vibration that consists of an out-of-phase combination of C_β–C_γ stretching and $N_\epsilon H_2$ rocking. This vibration also contains a significant C_δ–N_ϵ stretching component ($\sim$13%), which is in-phase with $N_\epsilon H_2$ rocking.

The in-phase combination of $N_\epsilon H_2$ rocking and C_δ–N_ϵ stretching of the $\sim$1110 cm^{-1} vibration is reminiscent of the AmIII mode of secondary amides. While complex, the AmIII vibration contains significant contributions of in-phase C–N stretching and N–H in-plane bending motions of the secondary amide group. We propose to call the $\sim$1110 cm^{-1} mode the Amide III^P (Am III^P) since the eigenvector composition of this vibration is analogous to that of the canonical AmIII of secondary amides. As discussed in detail below, the Am III^P vibration is sensitive to the χ_3 and χ_2 dihedral angles of Gln and Asn.

4.4.2 Conformational Dependence of the Am III^P Band

We performed DFT calculations on L-Gln molecules with χ_3 dihedral angles fixed at different values (see Computational Section for details) in order to identify spectroscopic markers that are diagnostic of the side chain χ_3 and χ_2 dihedral angles of Gln and Asn, respectively. We examined the frequency dependence of different primary amide vibrations and found that the Am III^P vibrational frequency and normal mode depends strongly on the OCCC dihedral angle.

Figure 4.3a shows the calculated cosinusoidal dependence of the Am III^P vibrational frequency on the χ_3 dihedral angle. The maximum frequency of the vibration occurs at χ_3 $\sim$0°, while minima occur near χ_3 $\pm$90°. The Gln Am III^P band frequency dependence on

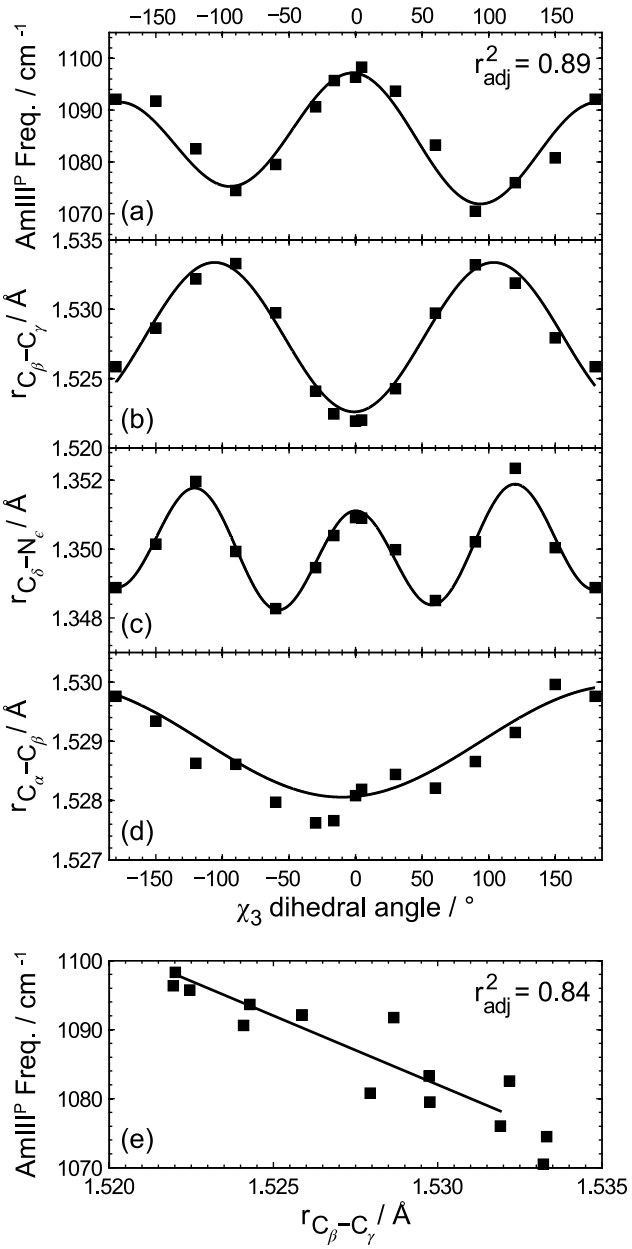


Figure 4.3: Calculated Am III$^{\mathrm{P}}$ frequency and bond length dependence on the χ_3 dihedral angle of the Gln side chain. (a) Am III$^{\mathrm{P}}$ frequency dependence; (b) C_β–C_γ bond length; (c) C_δ–N_ϵ bond length; (d) C_α–C_β bond length; and (e) shows the dependence of the Am III$^{\mathrm{P}}$ frequency on the C_β–C_γ bond length. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

the χ_3 angle follows a cosinusoidal relationship:

$$\nu(\chi_3) = \nu_0 + A\cos(2\chi_3) + B\cos(\chi_3 - C) \tag{4.1}$$

where $\nu_0 = 1084\,\mathrm{cm}^{-1}$, A = $10\,\mathrm{cm}^{-1}$, B = $3\,\mathrm{cm}^{-1}$, and C = $-31°$. These parameters were calculated from a least-squares fit of eq. 4.1 to the frequency dependence on the χ_3 angle in Figure 4.3a.

Figure 4.3a shows that the Am III^P frequency dependence on the χ_3 dihedral angle is asymmetric about χ_3 ~0°. This asymmetry is due to the chirality of L-Gln and L-Asn and leads to the requirement of two cosine terms to express the χ_3 frequency dependence of eq. 4.1. This is evident when we compare the L-Gln χ_3 dependence on the Am III^P frequency with that of butyramide (shown in Figure C1). In the case of butyramide, which is achiral, there is no asymmetry about 0°. As a result, the Am III^P frequency dependence on the OCCC dihedral angle of butyramide can be satisfactorily modeled with just one cosine term (eq. C.1).

4.4.3 Origin of the OCCC Dihedral Angle Dependence of the Am III^P Vibration

Understanding the conformational dependence of the Am III^P frequency on the primary amide OCCC dihedral angle requires a detailed knowledge of the atomic motions that give rise to the vibration. On the basis of our normal mode calculations of Gln, butyramide (Table C14), and propanamide [146], we conclude that $N_\epsilon H_2$ rocking, C_δ–N_ϵ stretching, and C_β–C_γ stretching define the Am III^P vibration. However, depending on the OCCC dihedral angle, other motions such as $C_\beta H_2$ twisting and C_α–C_β stretching can contribute to this vibration.

Therefore, we examined how the Gln C_δ–N_ϵ, C_β–C_γ, and C_α–C_β bond lengths change as a function of the χ_3 dihedral angle in order to understand the origin of the conformational sensitivity of the Am III^P vibration. Changes in these bond lengths impact the Am III^P frequency by affecting the vibrational mode bond force constants. As seen in Figure 4.3b-d, all the bond lengths show a dependence on the χ_3 dihedral angle. However, as seen in Figure 4.3b, the C_β–C_γ bond length shows the largest dependence on the χ_3 dihedral angle.

The Am III^P vibrational frequency has a strong correlation with the C_β–C_γ bond length, as shown in Figure 4.3e. The Am III^P vibrational frequency increases as the C_β–C_γ bond length decreases and *vice versa*.

The C_β–C_γ bond length dependence on the χ_3 dihedral angle appears to be due to hyperconjugation between the C_β–C_γ σ and the C_δ=O_ϵ π^* orbitals (Figure 4.4). This interaction is strongest when these orbitals maximally overlap, in the absence of significant phase cancellation due to the π^* orbital antisymmetry. When hyperconjugation occurs, the σ orbital donates electron density to the π^* orbital, which decreases the C_β–C_γ bond order and increases its bond length. This decreases the C_β–C_γ stretching force constant, which downshifts the Am III^P frequency.

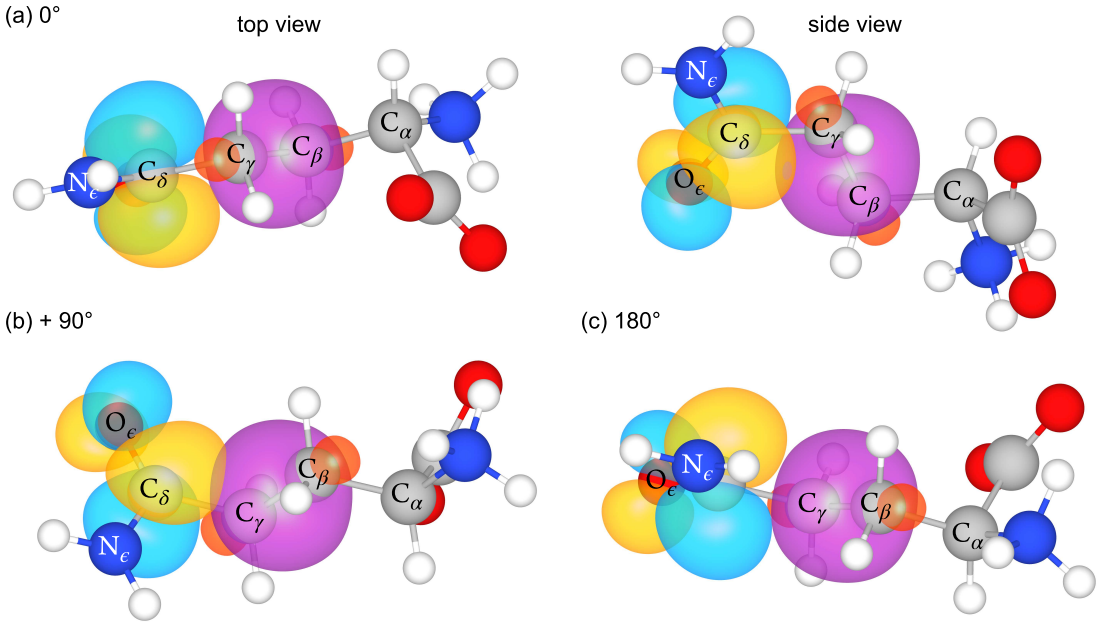


Figure 4.4: Hyperconjugation results in the C_β–C_γ bond length sensitivity to the χ_3 dihedral angle. Overlap of C_β–C_γ σ and C_δ=O_ϵ π^* NBO molecular orbitals when the χ_3 dihedral angle is: (a) 0°; (b) +90°; and (c) ±180°. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

We tested this hypothesis with natural bond orbital (NBO) analysis, which allows the DFT calculated electron densities to be displayed in terms of approximate σ and π^* molecular orbitals. According to our hypothesis, the C_β–C_γ bond length should be largest when

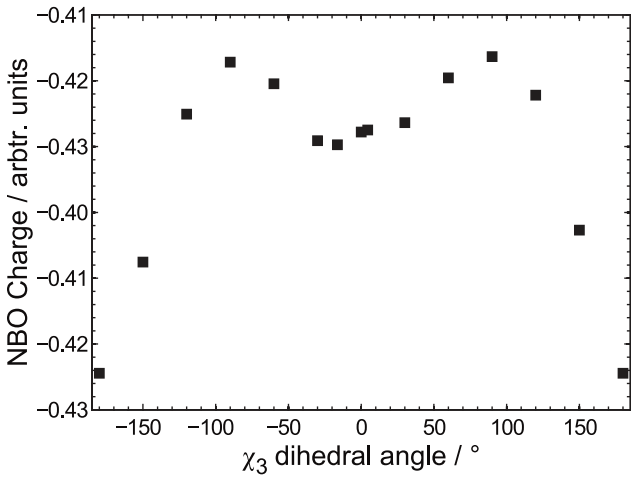


Figure 4.5: NBO charge of C_β in L-Gln as a function of the χ_3 dihedral angle. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

hyperconjugation is maximized and smallest when there is no hyperconjugation. Indeed, as seen in Figure 4.4b, there is significant overlap of the C_β–C_γ σ and C_δ=O_ϵ π^* NBO molecular orbitals at ±90°, where the C_β–C_γ bond length is largest. In contrast, at χ_3 ∼0°, where the C_β–C_γ bond length is shortest, the orbital overlap cancels due to the antisymmetry of the π^* orbital. Figure 4.5 shows the NBO charge on the C_β atom. As expected from our hyperconjugation hypothesis, the NBO C_β atom charge is less negative at $\chi_3 \sim \pm 90°$ compared to $\chi_3 \sim 0°$. The NBO C_β atom charge becomes even more negative at $\chi_3 \sim \pm 150°$ and $\chi_3 \sim \pm 180°$, even without additional hyperconjugation of the C_β–C_γ σ and C_δ=O_ϵ π^* orbitals. This result is likely an artifact because these extreme χ_3 dihedral angles are associated with physically impossible high energy structures that will be subject to other electron density alterations.

Our model accounts for the Am III^P frequency downshift as the dihedral angles approach $\chi_3 \sim \pm 90°$, where hyperconjugation is strongest. This behavior is the reverse of the Bohlmann effect [183–186], where a "negative" hyperconjugation transfers electron density from a lone pair orbital to an optimally positioned C–H σ^* orbital. This decreases the C–H bond order and substantially downshifts the C–H stretching frequencies.

4.4.4 Experimental Dependence of Am III^{P} Band Frequency on OCCC Dihedral Angle

We experimentally examined the dependence of the Am III^{P} band frequency on the primary amide OCCC dihedral angle by measuring the UVRR and visible Raman spectra of different Gln and Asn derivatives in the solid-state. We determined the structures of each of the different Gln and Asn derivative crystals with X-ray diffraction, and assigned the Am III^{P} band by performing DFT calculations and examining band shifts upon N-deuteration. Our X-ray diffraction methods and the band assignments of the crystals are discussed, in detail, in APPENDIX C.

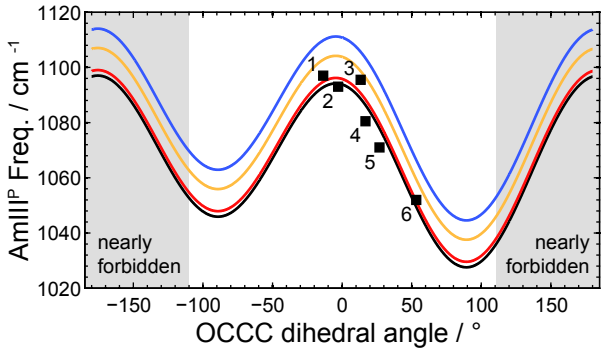


Figure 4.6: Experimental correlation of the Am III^{P} frequency to the χ_3 dihedral angle. The average frequency (from the 633 nm and 229 nm Raman spectra) of the Am III^{P} band was plotted as a function of the OCCC dihedral angle. 1 = L-Gln; 2 = Gly-Gln; 3 = D-Gln; 4 = GlnTBE; 5 = NAcGln; and 6 = Ser-Asn. The data were fit with eq. 4.2 (black line, r^2_{adj}= 0.83). The blue curve corresponds to eq. 4.3. The red curve corresponds to eq. 4.4. The yellow curve corresponds to eq. 4.5 and is an average of the red and blue curves. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

4.4.4.1 Dependence of Am III^{P} Band Frequency in Crystals

Figure 4.6 shows the Am III^{P} frequency dependence on the experimentally determined primary amide OCCC dihedral angles. We fit the experimental data to a function of the same form as eq. 4.1,

obtaining the following relationship:

$$\nu(\chi_3) = 1066\ (\mathrm{cm}^{-1}) + 29\ (\mathrm{cm}^{-1})\ \cos(2\chi_3) + 9\ (\mathrm{cm}^{-1})\ \cos(\chi_3 + 99^\circ) \tag{4.2}$$

which is shown in the Figure 4.6 black curve. To obtain the eq. 4.2 parameters, we fixed the A/B ratio to ∼3 as found in eq. 4.1, and performed a least squares minimization of the experimental data. Eq. 4.2 provides an excellent fit of the experimental data and captures the chiral asymmetry that occurs near $\chi_3 \sim \pm 90^\circ$.

4.4.4.2 Dependence of Am III^P Band Frequency for Fully Hydrated Primary Amides

The Am III^P band frequency also depends on the local hydrogen bonding and dielectric environment of the primary amide group [146]. In water, the Am III^P band of L-Gln is located at $\sim$1110 cm^{-1}, as compared with $\sim$1097 cm^{-1} in the solid-state. Based on Rhys et al.'s neutron diffraction study [187], the solution-state equilibrium structure of L-Gln in water does not appear to differ significantly from the single known L-Gln crystal structure [188]. From their solution-state structure, we determine that the equilibrium χ_3 dihedral angle of L-Gln in water is $\sim -12.8^\circ$. This differs by less than a degree (-13.54°) from the L-Gln crystal examined in this study. Thus, by setting the Am III^P frequency to 1110 cm^{-1}, χ_3 to -13.54°, and solving for ν_0, we obtain eq. 4.3

$$\nu(\chi_3) = 1083\ (\mathrm{cm}^{-1}) + 29\ (\mathrm{cm}^{-1})\ \cos(2\chi_3) + 9\ (\mathrm{cm}^{-1})\ \cos(\chi_3 + 99^\circ) \tag{4.3}$$

which is shown by the Figure 4.6 blue curve. This equation correlates the Am III^P band frequency to OCCC dihedral angles for situations in which the primary amide group is fully exposed to water, such as in PPII-like structures, 2.5_1-helices [189], and extended β-strand-like peptide conformations dissolved in water.

4.4.4.3 Dependence of Am III^P Band Frequency for Low Dielectric Constant and Weak Hydrogen Bonding Environments

The Am III^P frequency downshifts $\sim$15 cm^{-1} in the low dielectric and hydrogen bonding environment of acetonitrile compared to that in water (see APPENDIX C and Figure C6). This downshift derives from the different water versus acetonitrile stabilizations of the ground state $O_\epsilon{=}C_\delta N_\epsilon H_2$ and $^-O_\epsilon C_\delta{=}N_\epsilon H_2^+$

resonance structures of the primary amide group [146]. In both solvents, the $O_\epsilon{=}C_\delta N_\epsilon H_2$ resonance structure dominates; however, in acetonitrile the $^{-}O_\epsilon C_\delta{=}N_\epsilon H_2^+$ resonance structure contributes less than in water. Thus, the C_δ–N_ϵ bond length is larger in acetonitrile compared to water due to the lesser favorability of the $^{-}O_\epsilon C_\delta{=}N_\epsilon H_2^+$ resonance structure. Consequently, there is a smaller C_δ–N_ϵ stretching force constant in acetonitrile compared to water, which results in a downshift of the Am III^P frequency.

Eq. 4.3 can be modified in order to account for situations where the primary amide group is not engaged in significant hydrogen bonding interactions or when located in a low dielectric environment. We apply a $15\,cm^{-1}$ downshift in ν_0 from eq. 4.3 to determine eq. 4.4:

$$\nu(\chi_3) = 1068\ (cm^{-1}) + 29\ (cm^{-1})\ \cos(2\chi_3) + 9\ (cm^{-1})\ \cos(\chi_3 + 99^\circ) \tag{4.4}$$

which is shown in red in Figure 4.6.

4.4.4.4 Dependence of Am III^P Band Frequency for Unknown Dielectric and Hydrogen Bonding Environments

We suggest the use of eq. 4.5, which is the average of eqs. 4.3 and 4.4, for cases where the hydrogen bonding and dielectric environment of the primary amide group is unknown:

$$\nu(\chi_3) = 1076\ (cm^{-1}) + 29\ (cm^{-1})\ \cos(2\chi_3) + 9\ (cm^{-1})\ \cos(\chi_3 + 99^\circ) \tag{4.5}$$

It can be applied, for example, to determine the side chain χ_3 and χ_2 dihedral angles of Gln and Asn residues located in turn structures of proteins. For these residues, it may not be clear if the side chains are hydrogen bonded to water, to other side chains, or the peptide backbone. Eq. 4.5 is shown by the yellow curve in Figure 4.6.

4.4.5 Predicting Side Chain χ_3 and χ_2 Dihedral Angles in Gln and Asn as a Function of Ramachandran (Φ, Ψ) Angles

Shapovalov and Dunbrack [179] recently developed a new peptide backbone dependent rotamer library, which includes the non-rotameric Gln and Asn side chain χ_3 and χ_2 dihedral angles. Their database was compiled by analyzing high resolution crystal structures from

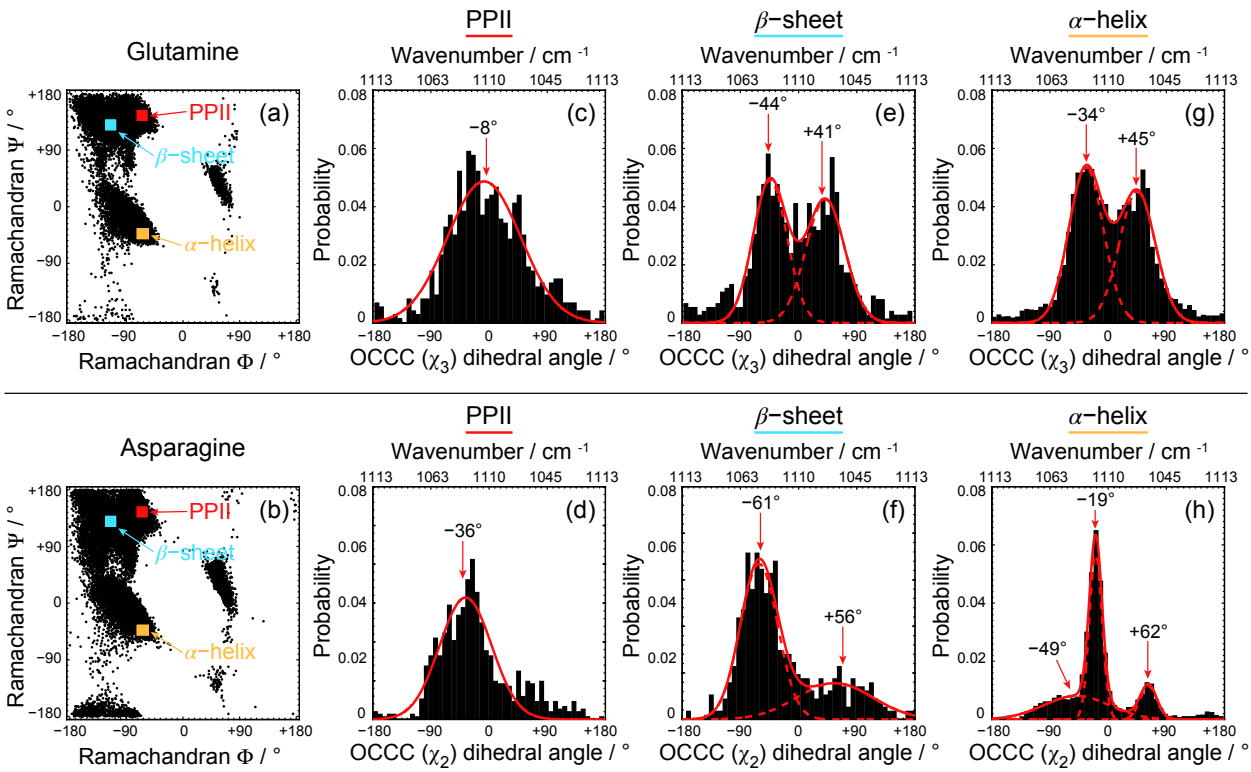


Figure 4.7: Gln and Asn side chain χ_3 and χ_2 dihedral angle dependence on secondary structure. Plots showing Ramachandran angles for PDB entries from the Shapovalov and Dunbrack database of (a) Gln and (b) Asn. The colored boxes correspond to canonical PPII (Φ= -65°, Ψ = 145°) angles (red), β-sheet (Φ = -115°, Ψ = 130°) angles (blue), and α-helix (Φ = -63°, Ψ = -43°) angles (cyan). Distributions of χ_3 and χ_2 dihedral angles for Gln and Asn residues that have (Φ, Ψ) angles close to canonical: (c, d) PPII-like, (e, f) β-sheet, and (g, h) α-helical structures. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

the protein data bank (PDB), and consists of ∼30,000 entries for Asn and ∼20,000 entries for Gln. Figure 4.7a and Figure 4.7b show Ramachandran plots of all of the Gln and Asn entries in the Shapovalov and Dunbrack database. The Gln and Asn side chains populate similar regions of the Ramachandran plot, and both show a preference for α-helical region (Φ, Ψ) angles. Asn populates a much broader range of (Φ, Ψ) angles, especially in the nearly forbidden "bridge" region between β-sheet and α-helical regions of the Ramachandran plot.

We used the Shapovalov and Dunbrack database to examine the side chain χ_3 and χ_2 dihedral angle preferences of Gln and Asn residues that possess canonical PPII, β-sheet, or α-helix Ramachandran angle values. Based on work by Richardson [190] and Karplus [191], we assume (Φ, Ψ) angles centered around ($-65°$, $145°$) for canonical PPII structures, ($-115°$, $130°$) for canonical β-sheets, and ($-63°$, $-43°$) for canonical α-helices. Figure 4.7c-h depict histograms of the χ_3 and χ_2 dihedral angles observed for the population of Gln and Asn residues with canonical PPII, β-sheet, or α-helical Ramachandran angles.

The Gln and Asn side chain χ_3 and χ_2 dihedral angles clearly depend upon the peptide bond Φ and Ψ angles. This correlation could result from a preference for particular χ_3 or χ_2 dihedral angles for stretches of consecutive peptide bonds with (Φ, Ψ) angles that result in PPII, β-sheet, or α-helical secondary structures. Alternatively, it could result from a preference for χ_3 or χ_2 dihedral angles for the (Φ, Ψ) angle values of their individual peptide bonds.

The χ_3 and χ_2 dihedral angle histograms of Gln and Asn residues that populate the canonical PPII region of the Ramachandran plot are shown in Figure 4.7c and Figure 4.7d. The distribution of χ_3 angles adopted by Gln is broader than that of the χ_2 angles of Asn. Both histograms are centered about negative dihedral angles, with Gln showing a peak at around $\chi_3 \sim -8°$ and Asn showing a peak near $\chi_2 \sim -36°$. It should be noted that the bias due to the L- amino acid chirality gives rise to a clear preference for negative χ_2 dihedral angles for the shorter side chain Asn residues.

The χ_3 and χ_2 dihedral angle histograms of Gln and Asn with β-sheet (Φ, Ψ) angles in Figure 4.7e and Figure 4.7f differ dramatically from one another. The population of Gln χ_3 dihedral angles (Figure 4.7e) is nearly symmetric about $\chi_3 \sim 0°$. The histogram is bimodal, with two peaks located near χ_3 angles of $\sim -44°$ and $\sim 41°$. In contrast, the population of Asn residues (Figure 4.7f) predominately adopts negative dihedral angles and is peaked around $\chi_2 \sim -61°$. A minor peak also occurs around $\chi_2 \sim 56°$.

Figure 4.7g and Figure 4.7h show histograms of the χ_3 and χ_2 dihedral angles of Gln and Asn residues that adopt canonical α-helical Ramachandran angles. As in Figure 4.7e, the Figure 4.7g Gln χ_3 dihedral angle population is roughly bimodal and nearly symmetric about $\chi_3 \sim 0°$. It is peaked at χ_3 angles of $\sim -34°$ and $\sim 45°$. In contrast, in Figure 4.7h the

population of Asn χ_2 dihedral angles is narrow and sharply peaked at $\chi_2 \sim -19°$ with two minor peaks at $\chi_2 \sim -49°$ and $\sim 62°$.

The χ_3 and χ_2 dihedral angle dependencies on the peptide bond Ramachandran angles, shown by the Shapovalov and Dunbrack database, enable us to predict the most probable Am III^P frequencies of Gln and Asn residues that adopt canonical PPII, β-sheet, and α-helix (Φ, Ψ) angles (shown in Table 4.3). For example, using eq. 4.3, we calculate that Gln and Asn side chains with PPII (Φ, Ψ) angles will have a maximum probability of showing Am III^P bands centered at $\sim 1111\,cm^{-1}$ and $\sim 1096\,cm^{-1}$, respectively. Similarly, we calculate that the Am III^P bands of Gln residues with β-sheet Ramachandran angles will have the greatest probability of being located at $\sim 1080\,cm^{-1}$ and/or $\sim 1089\,cm^{-1}$. In contrast, the Am III^P bands for Asn residues with β-sheet (Φ, Ψ) angles will have the largest probability of being located at $\sim 1064\,cm^{-1}$ and/or $\sim 1075\,cm^{-1}$. For α-helical Ramachandran angles, we calculate that the probability maxima for Am III^P bands will be at $\sim 1076\,cm^{-1}$ and/or $\sim 1098\,cm^{-1}$ for Gln and $\sim 1058\,cm^{-1}$, $\sim 1085\,cm^{-1}$, and/or $\sim 1107\,cm^{-1}$ for Asn residues.

Table 4.3: Predicted Am III^P Frequencies and OCCC Dihedral Angles for Gln and Asn Residues with Different Ramachandran Angles. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

			Gln		Asn	
	Φ (°)	Ψ (°)	χ_3 (°)	Am III^P Freq. (cm^{-1})	χ_2 (°)	Am III^P Freq. (cm^{-1})
PPII	-65	145	-8 (-22, -32)[a]	1111 (1106, 1099)[a]	-36	1096
β-sheet	115	130	-44, 41	1089, 1080	-6, 56	1075, 1064
α-helix	-63	-43	-34, 45	1098, 1076	-49, -19, 62	1085, 1107, 1058

[a]Values in parentheses were measured experimentally for Q_3 and $D_2Q_{10}K_2$.

We can calculate the expected Raman spectral Am III^P band shapes from the Gln χ_3 and Asn χ_2 dihedral angle histograms in Figure 4.7 using the Am III^P Raman band frequency dependencies of eqs. 4.2–4.5. These calculated band shapes (not shown) are unphysically broad ($>100\,cm^{-1}$). This clearly indicates that these histograms derive from the inhomogeneous distribution of χ_3 and χ_2 angles of individual Gln and Asn residues within the proteins found in the Shapovalov and Dunbrack database. This distribution of Raman frequencies from the calculated Am III^P band is much broader than the homogeneous linewidth of an Am III^P band expected for a single Gln and Asn residue in a typical PPII, β-sheet, or α-

helix conformation in proteins. The large widths of the Gln χ_3 and Asn χ_2 dihedral angle histograms result because the residues in the Shapovalov and Dunbrack database exist in a larger distribution of conformations, hydrogen bonding states, and chemical environments than we have so far encountered in our UVRR investigations.

4.4.6 Experimentally Determined Gln PPII-like Structure Peptide χ_3 Dihedral Angles

4.4.6.1 UVRR Spectra of Gln Peptides in PPII-like Structures We examined the UVRR spectra of two peptides, Q_3 and $D_2Q_{10}K_2$, in order to determine their solution-state χ_3 angles. Xiong et al. [125] previously showed that $D_2Q_{10}K_2$ exists in predominately PPII-like and 2.5_1-helix-like conformations when prepared using a "disaggregation" protocol developed by Wetzel and coworkers [77]. In this protocol, the $D_2Q_{10}K_2$ peptide is initially dissolved in a mixture of trifluoroacetic acid and hexafluoroisopropanol. These solvents are subsequently evaporated under dry N_2 gas, and the peptide is redissolved in pure water.

The UVRR spectra indicate that Q_3 has predominately PPII-like peptide bond conformations. Figure 4.8a shows the peak fitted ~204 nm excitation UVRR spectrum of Q_3 in the region between 1050 cm^{-1} to 1500 cm^{-1}. The $AmIII_3$ region, between ~1200 cm^{-1} to 1280 cm^{-1}, is most sensitive to the secondary structure of the peptide since its frequency depends on the Ramachandran Ψ angle [96, 97]. This region is well fit by two Gaussian bands located at ~1210 cm^{-1} and ~1260 cm^{-1}. Using the methodology of Mikhonin *et al.* [97], we correlated the band peak positions to their Ψ angles. We used their eq. 6A to correlate the 1210 cm^{-1} frequency of the $AmIII_3$ band to a Ψ angle of 103°±3° and the 1260 cm^{-1} frequency to a Ψ angle of 157°±2°. The Ψ angle of ~157° derives from peptide bonds situated in PPII-like conformations, while the Ψ angle of ~103° derives from peptide bond situated in β-strand-like conformations. Assuming identical Raman cross sections for these two different species, we find that the peptide bonds are dominated by PPII-like Ψ angles (~87±2%), while a small fraction adopt β-strand-like Ψ angles (~13±2%). This is supported by the circular dichroism spectra of Q_3 shown in Figure C7, which show a predominantly PPII spectral signature.

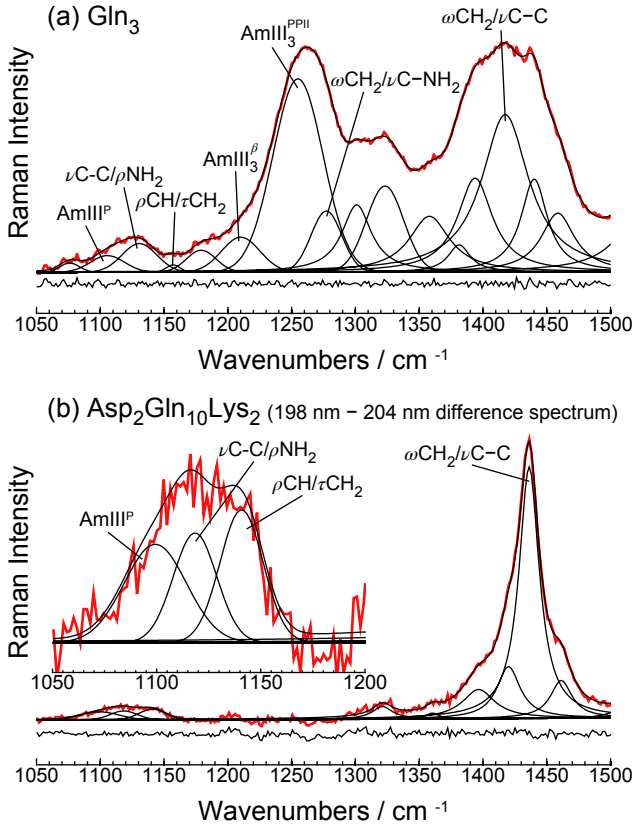


Figure 4.8: Deconvolution of the UVRR spectra of Q_3 and $D_2Q_{10}K_2$. (a) Fitting the 204 nm excitation UVRR spectrum of Q_3. (b) 198-204 nm difference spectrum of $D_2Q_{10}K_2$ taken from Xiong *et al.* [125] The inset shows the Am III^P region of $D_2Q_{10}K_2$. The χ^2_{red} statistics for the spectral fits shown in (a) and (b) are 1.1 and 0.74, respectively. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

4.4.6.2 χ_3 Dihedral Angle Determination of Side Chains in Gln Peptides

The Am III^P bands of Q_3 and $D_2Q_{10}K_2$ are found in the region between ~1050 cm^{-1} and 1150 cm^{-1}. Based on our normal mode analysis of Gln, we fit this region with four bands that derive from C_αH rocking/$C_\gamma H_2$ twisting, C_β–C_γ stretching/$N_\epsilon H_2$ rocking, NH_3 rocking/C_α–N stretching, and the Am III^P vibrations. For Q_3, these bands are located within a broad asymmetric spectral feature at ~1080 cm^{-1}, ~1106 cm^{-1}, ~1130 cm^{-1}, and ~1160 cm^{-1}.

We assigned these bands based on our analysis of Gln. We assign the ~1160 cm^{-1} band to a C_αH rocking/$C_\gamma H_2$ twisting mode, the ~1130 cm^{-1} band to a C_β–C_γ stretching/$N_\epsilon H_2$ rocking vibration, and the ~1080 cm^{-1} band to a NH_3 rocking/C_α–N stretching mode. The ~1106 cm^{-1} band appears as a low-frequency shoulder feature and is assigned to the Am III^P vibration. This is very close to the predicted Am III^P vibrational frequency band center from the Gaussian fit of PPII-like structures in Figure 4.7c, as listed in Table 4.3. In fact, the Am III^P frequency band center of Q_3 differs by only ~5 cm^{-1} from the predicted frequency band center (~1111 cm^{-1}) for PPII Ramachandran angles.

Figure 4.8b shows the 198 nm – 204 nm difference spectrum of disaggregated $D_2Q_{10}K_2$ published by Xiong *et al.* [125] Xiong et al. showed that excitation at 198 nm enhances the primary amide UVRR bands more than does excitation at 204 nm. Thus, the Figure 4.8b $D_2Q_{10}K_2$ difference spectrum is dominated by the primary amide Gln side chain bands with little interference from the secondary amide peptide bond UVRR bands.

The inset in Figure 4.8b shows the region where the Am III^P band of $D_2Q_{10}K_2$ is located. We parsimoniously peak fit this region to three Gaussian bands located at ~1099 cm^{-1}, ~1118 cm^{-1}, and ~1140 cm^{-1}. Using prior knowledge from our analysis of Gln, we assign the bands at ~1118 cm^{-1} and ~1140 cm^{-1} to the the C_β–C_γ stretching/$N_\epsilon H_2$ rocking and C_αH rocking/$C_\gamma H_2$ twisting vibrations, respectively. The ~1099 cm^{-1} band is assigned to the Am III^P band.

The Am III^P bandwidths of Q_3 and $D_2Q_{10}K_2$ are ~30 cm^{-1}, which is similar to that of Gln in H_2O (Figure 4.2a). These bandwidths are roughly twice as large as those found in the Raman spectra of the different Gln and Asn derivative crystals, which we measure to be on average ~13.3(50) cm^{-1}. This bandwidth is significantly larger than our spectrometer resolution of ~4.5 cm^{-1}. Thus, if we assume a Lorentzian band shape, we estimate that the

Am III^{P} band homogeneous linewidth for a Gln compound with a well-defined χ_3 angle is $\sim$6.6 cm^{-1}. The fact that the Am III^{P} bandwidths of solution-state Gln, Q_3, and $D_2Q_{10}K_2$ are much broader than those measured in our crystals suggests that there is a distribution of hydrogen bonding states and χ_3 angles in these compounds.

Given the estimated homogeneous linewidth, we can roughly calculate the distribution of χ_3 angles of Gln, Q_3, and $D_2Q_{10}K_2$ by using a methodology that is similar to that of Asher *et al.* [105] To do this, we assume that the inhomogeneously broadened Am III^{P} bands derive from a distribution of different χ_3 dihedral angles, which can be represented as the sum of M Lorentzian bands:

$$A(\nu) = \frac{1}{\pi} \sum_{i=1}^{M} I_i \frac{\Gamma^2}{\Gamma^2 + (\nu - \nu_i)^2} \tag{4.6}$$

where I_i is the intensity of a Lorentzian band that occurs at a given center frequency, ν_i, and Γ is the homogeneous linewidth.

We can apply eq. 4.3 to correlate the ν_i Am III^{P} frequencies of the M Lorentzian bands to their corresponding χ_3 dihedral angles. As shown in Figure 4.6, a single Am III^{P} frequency can correspond to as many as four possible χ_3 dihedral angles. However, the Shapovalov and Dunbrack database show that χ_3 dihedral angles that are greater than +90° and less than -90° are nearly forbidden (Figure 4.7). Thus, we consider only the two χ_3 dihedral angle solutions that are found in the region between -90° and +90°, as shown in Figure 4.9.

To determine which of the two remaining χ_3 dihedral angle solutions is occurring in our peptides, we first fit the histograms to the sum of two Gaussians with identical amplitudes, A, and widths, w, but different center χ_3 angles, $\bar{\chi}_{3,1}$ and $\bar{\chi}_{3,2}$:

$$I(\chi_3) = Ae^{-(\frac{\chi_3 - \bar{\chi}_{3,1}}{w})^2} + Ae^{-(\frac{\chi_3 - \bar{\chi}_{3,2}}{w})^2} \tag{4.7}$$

The Gln, Q_3, and $D_2Q_{10}K_2$ results all show one Gaussian centered at negative χ_3 angles and another Gaussian centered at positive χ_3 angles (Figure 4.9). For Gln, we assume that the Gaussian centered at $\chi_3 \sim -13°$ is the physically relevant solution based on the neutron diffraction study of Rhys *et al.* [187]. For Q_3 and $D_2Q_{10}K_2$, we conclude that the Gaussians centered at negative χ_3 angles correspond to the physically relevant solutions to eq. 4.3

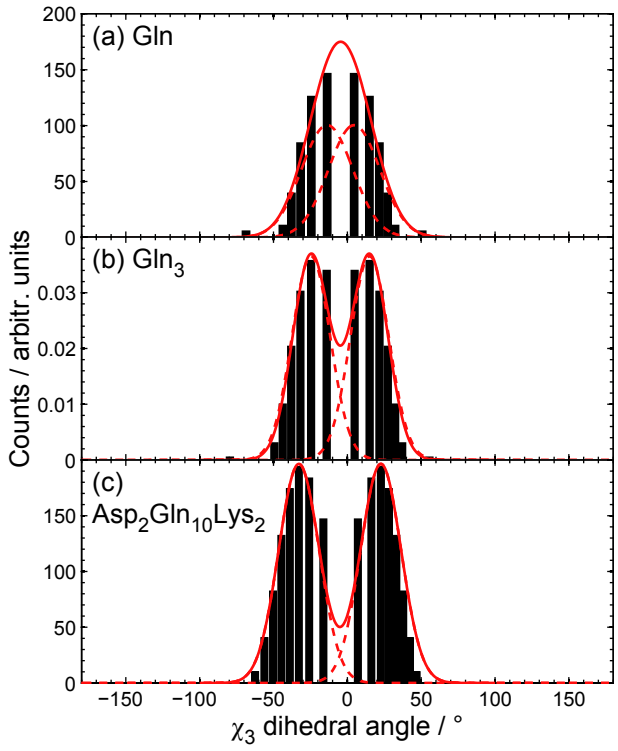


Figure 4.9: χ_3 dihedral angle historgrams calculated by decomposing Am III$^{\mathrm{P}}$ bands into a sum of Lorentzians for (a) Gln, (b) Q_3, and (c) $D_2Q_{10}K_2$ in water. Because the solution to eq. 4.3 is double valued between $\pm 90°$, the histograms show two peaks. The histograms were fit to two identical Gaussians that differed only in center χ_3 dihedral angles (shown in dashed lines). The sum of the Gaussians is shown in the solid red lines. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

because they fall within the range of χ_3 dihedral angles most commonly adopted by Gln residues that populate PPII (Φ, Ψ) angles (Figure 4.7c).

Figure 4.10 shows the resulting χ_3 dihedral angle distributions for Gln, Q_3, and $D_2Q_{10}K_2$ by assuming the physically relevant solutions to eq. 4.3. The distributions of Q_3 and $D_2Q_{10}K_2$ populate χ_3 angles similar to that of Gln. This suggests that primary amides of Q_3 and $D_2Q_{10}K_2$ are fully solvated like that of monomeric Gln in water. Thus, the Gln side chains are not engaged in side chain-backbone peptide bond hydrogen bonding as previously hypothesized [33].

4.4.6.3 Determination of the Gibbs Free Energy Landscape for Gln and Gln peptides Along the χ_3 Dihedral Angle Reaction Coordinate

The structure sensitivity of the Am III^P band enables us to determine the Gibbs free energy landscape of the Gln side chains along the χ_3 dihedral angle structure coordinate. To do this, we assume that the probability of each $\chi_{3,i}$ angle in the χ_3 dihedral angle distributions of Gln, Q_3, and $D_2Q_{10}K_2$ shown in Figure 4.10a-c is given by a Boltzmann distribution:

$$\frac{p(\chi_{3,i})}{p(\chi_{3,0})} = e^{-(\Delta G(\chi_{3,i})/RT)} \tag{4.8}$$

where $p(\chi_{3,i})/p(\chi_{3,0})$ is the ratio of populations with χ_3 angles $\chi_{3,i}$ and $\chi_{3,0}$. The angle, $\chi_{3,0}$, is the minimum energy χ_3 angle, R is the molar gas constant, T is the experimental temperature (293 K), and $\Delta G(\chi_{3,i}) = G(\chi_{3,i}) - G(\chi_{3,0})$. We assume in eq. 4.8 that each $\chi_{3,i}$ dihedral angle state has a degeneracy of one.

To calculate the free energy difference, $\Delta G(\chi_{3,i})$, between a particular $\chi_{3,i}$ angle and the equilibrium $\chi_{3,0}$ angle, we rearrange eq. 4.8:

$$\Delta G(\chi_{3,i}) = -RT \ln \left[\frac{p(\chi_{3,i})}{p(\chi_{3,0})}\right] \tag{4.9}$$

Figure 4.11 shows the calculated Gibbs free energy landscapes of Gln, Q_3, and $D_2Q_{10}K_2$ along the χ_3 dihedral angle structure coordinate. We model the side chain free energies

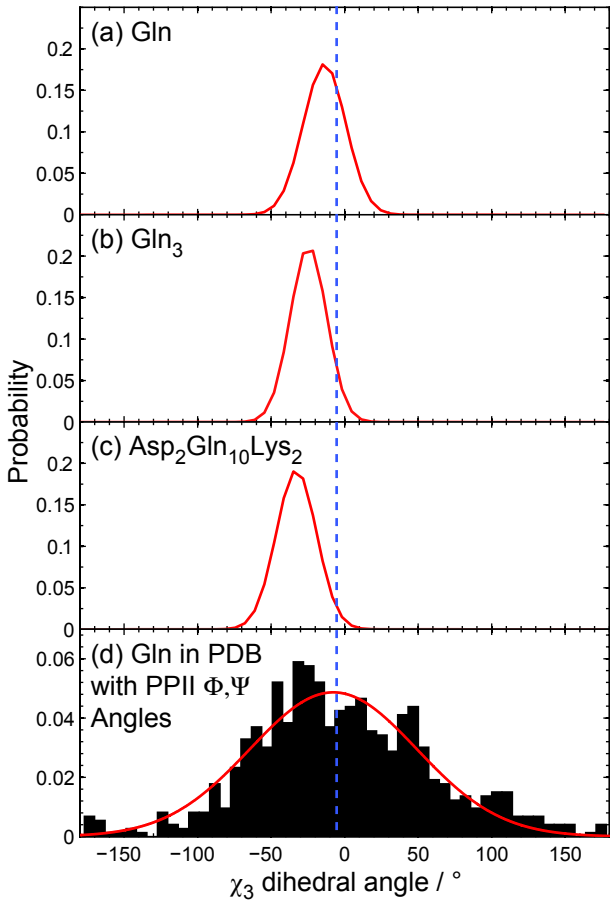


Figure 4.10: Comparison of χ_3 dihedral angle distributions between (a) Gln, (b) Q_3 in a predominately PPII-like conformation, and (c) $D_2Q_{10}K_2$ in a PPII/2.5_1-helix equilibrium. (d) χ_3 angle distribution of Gln residues with PPII-like Ramachandran angles from the Shapovalov and Dunbrack database. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

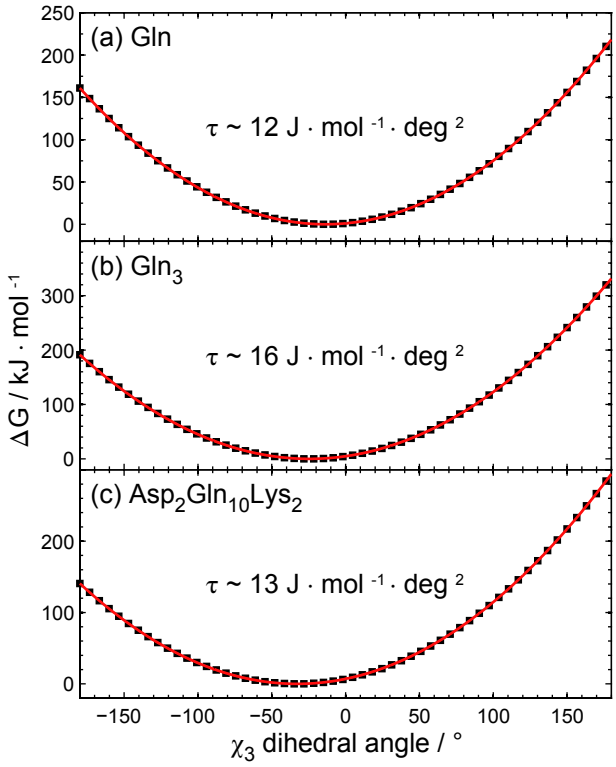


Figure 4.11: Gibbs free energy landscapes of (a) Gln, (b) Q_3, and (c) $D_2Q_{10}K_2$. The energy wells can be modeled by assuming a harmonic oscillator model torsional spring: $\Delta G(\chi_{3,i}) = \frac{\tau}{2}(\chi_{3,i} - \chi_{3,0})^2$, where τ is the torsional spring force constant and $(\chi_{3,i} - \chi_{3,0})$ is the displacement from the equilibrium position. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

about the equilibrium $\chi_{3,0}$ angles in terms of a simple Hooke's Law torsional model:

$$\Delta G(\chi_{3,i}) = \frac{\tau}{2}(\chi_{3,i} - \chi_{3,0})^2 \tag{4.10}$$

where τ is the torsional force constant. We can fit the free energy landscapes in Figure 4.11 to eq. 4.10 to determine the torsional force constants along the χ_3 dihedral angle coordinate of Gln, Q_3, and $D_2Q_{10}K_2$. We find that τ ~12 J·mol^{-1}·deg^{-2} for Gln, ~16 J·mol^{-1}·deg^{-2} for Q_3, and ~13 J·mol^{-1}·deg^{-2} for $D_2Q_{10}K_2$. The similarity of the Q_3 and $D_2Q_{10}K_2$ χ_3 angle torsional force constants of Gln most likely results from the similar side chain constraints and solvation states of these compounds.

4.5 CONCLUSIONS

We determined the dependence of the Am III^P band frequency on the χ_3 and χ_2 dihedral angles of Gln and Asn side chains. The Am III^P vibration is complex and consists of C_δ–N_ϵ (C_γ–N_δ) stretching and $N_\epsilon H_2$ ($N_\delta H_2$) rocking motions that are out-of-phase with C_β–C_γ (C_α–C_β) stretching in Gln (Asn). The frequency of the Am III^P vibration shows a cosinusoidal dependence on the χ_3 and χ_2 dihedral angles of the Gln and Asn side chains. The structural sensitivity of the Am III^P vibration derives from hyperconjugation between the C_β–C_γ (C_α–C_β) σ and the C_δ=O_ϵ (C_γ=O_δ) π^* orbitals. Hyperconjugation between these two orbitals increases the C_β–C_γ (C_α–C_β) bond length, which decreases the C_β–C_γ (C_α–C_β) stretching force constant and causes a downshift in the Am III^P frequency. In this case, hyperconjugation gives rise to spectroscopic markers diagnostic of local dihedral angles. This suggests that future studies of conformationally dependent hyperconjugation interactions will enable the discovery of new, structurally sensitive spectroscopic makers.

The correlations between the Am III^P frequency and the χ_3 and χ_2 dihedral angles of Gln and Asn side chains will be useful for protein conformational investigations, particularly for amyloid-like fibril and prion aggregates. In general, fibril and prion aggregates are insoluble and cannot be crystallized. Therefore, there are few approaches to obtain molecular-level structural information. As a result, little is known about the structure of Gln and Asn side chains in fibrils. The Am III^P spectroscopic marker band enables us to experimentally probe conformations of the Gln side chains of polyQ fibrils in order to obtain new, molecular-level insights into fibril structures.

4.6 ACKNOWLEDGMENT

Funding for this work was provided by the University of Pittsburgh. E.M.D. gratefully acknowledges support through the NIH Molecular Biophysics and Structural Biology Training Grant (T32 GM88119). The computational work was supported by the University of Pittsburgh Center for Simulation and Modeling through the super-computing resources pro-

vided. We thank Dr. Sergei V. Bykov for useful discussions, and Prof. Roland Dunbrack, Jr. for providing his rotamer database.

5.0 POLYGLUTAMINE FIBRILS: NEW INSIGHTS INTO ANTIPARALLEL β-SHEET CONFORMATIONAL PREFERENCE AND SIDE CHAIN STRUCTURE

Adapted with permission from: David Punihaole, Riley J. Workman, Zhenmin Hong, Jeffry D. Madura, Sanford A. Asher. Polyglutamine Fibrils: New Insights into Antiparallel β-sheet Conformational Preference and Side Chain Structure. *The Journal of Physical Chemistry B*, **2016**, 120, 3012–3026.

Author Contributions: D.P. acquired and analyzed the UVRR, electron microscopy, and X-ray diffraction data. R.J.W. performed molecular dynamics simulations and analyzed the data. Z.H. performed DFT calculations and analyzed the data with the assistance of D.P. The manuscript was prepared by D.P. and S.A.A. with the assistance of R.J.W. and J.D.M.

Understanding the structure of polyQ amyloid-like fibril aggregates is crucial to gaining insights into the etiology of at least ten neurodegenerative disorders, including Huntington's disease. Here, we determine the structure of $D_2Q_{10}K_2$ (Q10) fibrils using UVRR spectroscopy and Molecular Dynamics (MD). Using UVRR, we determine the fibril peptide backbone Ψ and Gln side chain χ_3 dihedral angles. We find that most of the fibril peptide bonds adopt antiparallel β-sheet conformations; however, a small population of peptide bonds exist in parallel β-sheet structures. Using MD, we simulate three different potential fibril structural models that consist of either β-strands or β-hairpins. Comparing the experimentally measured Ψ and χ_3 angle distributions to those obtained from the MD simulated models, we conclude that the basic structural motif of Q10 fibrils is an extended β-strand structure.

Importantly, we determine from our MD simulations that Q10 fibril antiparallel β-sheets are thermodynamically more stable than parallel β-sheets. This accounts for why polyQ fibrils preferentially adopt antiparallel β-sheet conformations, instead of in-register parallel β-sheets like most amyloidogenic peptides. In addition, we directly determine, for the first time, the structures of Gln side chains. Our structural data give new insights into the role that the Gln side chains play in the stabilization of polyQ fibrils. Finally, our work demonstrates the synergistic power and utility of combining UVRR measurements and MD modeling in order to determine the structure of amyloid-like fibrils.

5.1 INTRODUCTION

There are at least ten neurodegenerative disorders, including Huntington's disease, that are associated with mutational expansions in genomic CAG codon repeats [1]. These expansions increase the length of polyQ repeats in proteins. The increase in the repeat length of polyQ segments greatly enhances protein misfolding and aggregation. Although the exact mechanism of neurotoxicity is still heavily debated, the pathological hallmark of all CAG repeat diseases is the formation of large neuronal inclusions composed of polyQ-rich aggregates [67, 192, 193]. Given their potential role in neurotoxicity, it is therefore crucial to understand the structure of polyQ-rich aggregates.

Numerous structures have been proposed for polyQ fibrils based on the results of many different biophysical methods. For example, X-ray diffraction studies indicate that polyQ fibrils of various Gln repeat lengths all show similar X-ray diffraction patterns. Despite this, these studies assign very different structures from surprisingly similar X-ray data, including "polar-zippers" [52], β-helices [53], and canonical β-sheet structures [54, 55]. More recent structural studies of polyQ fibril aggregates use solid state NMR. One of these studies [56], concludes that the basic structural motif of polyQ fibrils prepared from pathologically relevant peptides is a "β-arc," similar to that of Aβ [57, 194]. The β-arc model has been challenged by other solid state NMR and biochemical studies [58, 59], which alternatively suggest that polyQ fibrils are composed of extended β-strands that contain reverse hairpin

turns.

In recent years, several Molecular Dynamics (MD) approaches have also been used to investigate the structural properties of polyQ-rich fibrils. The bulk of these computational studies focus mainly on the kinetic or thermodynamic stability of different fibril structures [51, 63, 66, 195]. However, these MD studies are conducted independently of experimental studies. Thus, there is little *direct* validation of the computational results against experimental data.

The lack of consensus regarding the structure of polyQ fibrils underscores the need for new and incisive biophysical methods that can *quantitatively* discriminate between the many proposed models. A fundamental factor in understanding polyQ fibrils is determining the structures and hydrogen bonding environments of the Gln side chains, which are thought to play an important role in stabilizing the aggregates. Another important structural property to understand is the propensity of polyQ fibrils to adopt antiparallel β-sheets instead of in-register parallel β-sheets, like most amyloidogenic peptides and prions.

UVRR spectroscopy is a powerful biophysical tool for studying the conformational ensembles and aggregation dynamics of amyloidogenic peptides [90–95]. An advantage of UVRR is that quantitative information can be obtained quickly, under dilute conditions, and without the need for extensive or complex isotopic labeling of peptides and proteins [89]. In recent years, numerous UVRR spectroscopic markers have been discovered. Some of these markers are sensitive to the Ramachandran Ψ angles of the peptide bonds [96, 97], while others are sensitive to the dihedral angles of amino acid side chains [169, 171, 173], including Asn and Gln [180]. Other marker bands are sensitive to the hydrogen bonding and the dielectric environments of peptide bonds and side chains [101, 103, 123, 124, 146, 171].

We can combine structural information obtained from interpreting these spectral markers with results from MD simulations to determine the structure of polyQ and other amyloid-like fibrils. An elegant example of this approach was recently published by Buchanan *et al.* [197], who combined 2D IR spectroscopy with MD simulations to determine the structure of $K_2Q_{24}K_2W$ fibrils. They concluded that $K_2Q_{24}K_2W$ fibrils adopt an antiparallel β-sheet structure that contains β-turns, but not β-arc structures, after comparing their experimental Amide I spectra with those calculated from simulated models.

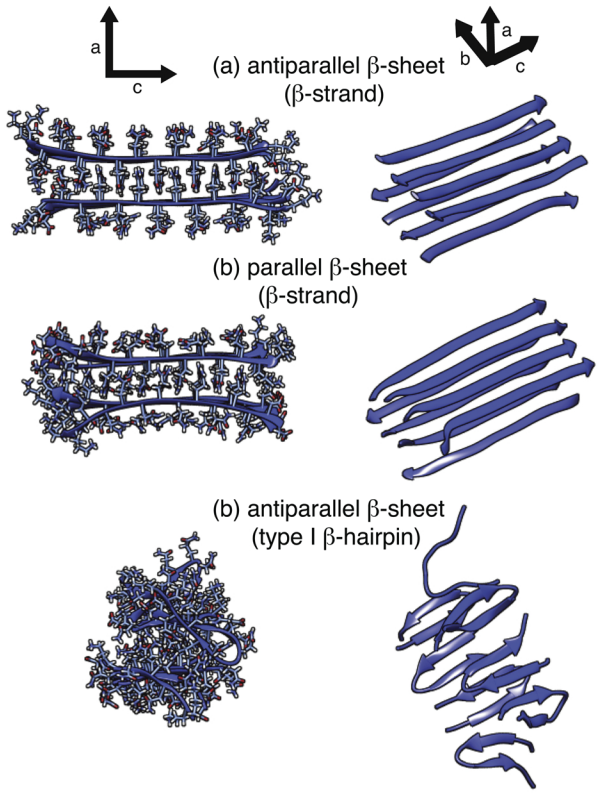


Figure 5.1: MD simulated structures of model Q10 fibril systems in (a) a β-strand configuration with an antiparallel β-sheet architecture, (b) a β-strand configuration with a parallel β-sheet architecture, and (c) a Type I β-hairpin configuration with an antiparallel β-sheet architecture. The a,b, and c axes refer to the inter-sheet, inter-strand, and intra-chain dimensions, respectively, of the fibril models. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

In this work, we synergistically couple UVRR and MD in order to determine the structures of polyQ amyloid-like fibrils prepared from the model peptide $D_2Q_{10}K_2$ (Q10). Xiong *et al.* [125] previously showed that this peptide can exist in two distinct solution-state conformations, a putative β-hairpin-like structure (called NDQ10) and a PPII-like structure (called DQ10). We show that both NDQ10 and DQ10 peptide solutions can aggregate into amyloid-like fibrils. We use UVRR to measure the Ramachandran Ψ angle distributions of the NDQ10 and DQ10 fibril peptide bonds, as well as their Gln χ_3 ($O_{\epsilon 1}$–C_δ–C_γ–C_β) side chain dihedral angles.

To determine the structure of NDQ10 and DQ10 fibrils, we compare our experimentally determined Ψ and χ_3 angle distributions to those obtained from three MD simulated fibril models (Figure 5.1). On the basis of these comparisons, we find that Q10 fibrils consist of extended β-strands that predominately assemble into antiparallel β-sheets; however, small populations of Q10 peptide bonds are in parallel β-sheet structures. From our MD simulations, we determine that polyQ fibrils in antiparallel β-sheets are lower in free energy compared to parallel β-sheets. This energetic preference appears unique for polyQ fibrils compared to typical amyloid-like fibrils, where in-register parallel β-sheets are thought to be at lower energy [198]. Finally, our χ_3 dihedral angle measurements and MD simulations of Q10 fibrils leads us to propose a new model for the structure of Gln side chains in polyQ fibrils.

5.2 EXPERIMENTAL SECTION

5.2.1 Materials

The 14 residue peptide, $D_2Q_{10}K_2$ (Q10), was purchased from AnaSpec Inc. at $\geq$95% purity. Trifluoroacetic acid (TFA) was purchased from Acros at 99.5% purity, and 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP) was obtained from Fluka at $\geq$99% purity. HPLC-grade H_2O was purchased from Fisher Scientific, and D_2O (99.9 atom % D) was purchased from Cambridge Isotope Laboratories, Inc. NaOD (40 wt % solution in D_2O, 99+ atom % D) and L-glutamine (99% purity) was purchased from Sigma Aldrich.

5.2.2 Sample Preparation

NDQ10 peptide solutions were prepared by dissolving Q10 directly in H_2O or D_2O. DQ10 solutions were prepared using a standard protocol based on a method developed by Wetzel and coworkers [77]. Briefly, DQ10 samples were prepared by suspending the lyophilized Q10 peptide powder received from Anaspec Inc. in a 1:1 (v/v) mixture of TFA and HFIP. The samples were sonicated for 20 min and incubated at room temperature for 2 h. The solvents

were evaporated under a gentle stream of dry N_2 gas for 1 h. The peptide film was dissolved in H_2O or D_2O and ultracentrifuged at 627,000 × g for 30 min. at 4 °C. The top 2/3 of the solution was decanted and used for the aggregation reaction.

Fibril aggregates were prepared by initially dissolving NDQ10 and DQ10 prepared peptides in H_2O or D_2O at 5 mg mL^{-1} concentrations and incubating the samples at 60 °C in vials sealed with Teflon tape. Solutions were titrated to pH 7 (pD 8) using NaOH (NaOD) solutions. After incubation for 6 days, aggregates were harvested via centrifugation, and the pellets resuspended in 120 µL of H_2O or D_2O. For the hydrogen-deuterium exchange (HX) experiments, harvested aggregates prepared in H_2O (D_2O) were washed in 500 µL D_2O (H_2O), centrifuged, and the supernatant removed. The pellets were resuspended in 120 µL of D_2O (H_2O) and incubated for 3 h at room temperature in sealed vials.

5.2.3 Transmission Electron Microscopy (TEM)

A 10 µL aliquot of aggregate solution was placed onto a freshly glow discharged carbon-coated grid for 2 min before blotting dry with filter paper. Samples were stained with 10 µL of 1% (w/v) uranyl acetate for 2 min, and the excess stain removed by blotting the grid. Grids were imaged using a Tenai F20 electron microscope (FEI Co.) operating at 200 kV and equipped with a 4k × 4k CCD camera (Gatan).

5.2.4 X-ray Diffraction of Fibril Films

Aggregates prepared from the NDQ10 peptide solutions were placed into the wide end of a 0.7 mm diameter quartz X-ray capillary tube (Charles Supper Company). The wide end of the tube was sealed with melted beeswax and the thin end of the tube left open for drying. DQ10 aggregates were mounted on the broken end of a quartz capillary tube and dried. Aggregate samples were placed on the end of a goniometer head, and centered in the X-ray beam path. X-ray diffraction patterns were collected using a Bruker X8 Prospector Ultra with a Copper micro-focus tube ($\lambda = 1.541\,78$ Å) and equipped with an Apex II CCD detector. All data were collected at room temperature with exposure times of 60 min.

5.2.5 UVRR spectroscopy

The UVRR instrumentation used is described in detail by Bykov *et al.* [131] Briefly, ∼204 nm light was obtained by mixing the third harmonic with the 816 nm fundamental generated by a tunable Ti:sapphire system (Photonics Industries) operating at 1 kHz. An Indigo S tunable Ti:sapphire system (Positive Light), operating at 5 kHz, generated ∼197 nm light by mixing the third harmonic with the ∼788 nm fundamental. For fibrils measured in solution, the laser light was focused onto a spinning Suprasil quartz NMR tube containing the sample. The average laser power at the sample ranged from ∼0.4 - 0.5 mW. A ∼165° backscattering geometry was used. The total acquisition time to collect spectra was only ∼10 min. The scattered light was imaged into a home-built subtractive double monochromator and detected with a liquid N_2 cooled, back-thinned Spec-10:400B CCD camera (Princeton Instruments) with a Lumogen E coating. The spectrometer resolution was ∼5 cm^{-1} at the excitation wavelengths used. A description of how the spectra were processed is in APPENDIX D.

5.3 COMPUTATIONAL SECTION

We considered three model fibril systems, as shown in Figure 5.1. These systems are composed of eight Q10 peptides that are assembled into parallel and antiparallel β-sheets. The two β-sheets were oriented parallel to each other, but rotated by 180°, in order to maximize attractive electrostatic interactions between terminal Asp and Lys residues. Models *a* and *b* were constructed using canonical β-sheet Ramachandran dihedral angles [199]. In the case of model *a*, we used canonical antiparallel β-sheet (Φ, Ψ) angles of (−140°, 135°). For model *b*, we used canonical parallel β-sheet (Φ, Ψ) angles of (−120°, 113°). Model *c* was constructed using β-hairpin geometries observed in metadynamics simulations (data not shown). These model fibrils were constructed using the Molecular Operating Environment (MOE 2013.10) software suite [199] and were solvated in a water box. Files containing the initial structure coordinates used in the MD simulations, as well as the NAMD configuration

templates, are available for download.

After constructing the models, the solvated fibril systems were energy minimized for 10,000 steps using the conjugate gradient method and then equilibrated for 50 ps. During equilibration, the fibril atoms were initially restrained by harmonic potentials. After relaxing the water around the restrained fibrils, the water molecules and peptide side chains were then energy-minimized for 10,000 steps. This was followed by 100 ps of equilibration, in which the peptide backbone atoms were restrained. The model fibril systems were then equilibrated for an additional 50 ns without restraints. All fibril models retained structural integrity throughout the energy minimization and equilibration. The fibril models were simulated using classical MD for 200 ns.

MD simulations were performed with the NAMD software package (version 2.10) [200]. The potential energies were calculated with the CHARMM22/CMAP force field [201]. This force field was chosen for its torsional energy corrections intended to decrease α-helix bias and stabilize β-strand secondary structures. Other force fields, such as Amber99ffsb [202], also implement these corrections; however, CHARMM22/CMAP [201] was selected due to its reported accuracy in describing α-helix, β-sheet, and disordered structures [203]. The fibril models were solvated using the solvate module of VMD 1.9.1 [204], resulting in a periodic box of 5,000 water molecules with dimensions of $70 \times 50 \times 50\,\text{Å}^3$ for a total system size of 17,189 atoms. The TIP3P water model [205] was employed in all simulations, and the particle mesh Ewald algorithm [206] was used with a grid spacing of 1.0 Å to calculate full system electrostatics. An integration time step of 2 fs was employed. Simulations were performed under the NPT ensemble, with a Langevin thermostat and piston utilized to regulate the temperature of 300 K and pressure of 1.013 25 bar, respectively. The pair interaction cutoff was 12.0 Å, and the switch distance was 10.0 Å.

Simulation analysis was performed using VMD 1.9.1 and Tcl scripting. The Ψ and χ_3 dihedral angles were obtained for all Gln residues with a Tcl script in VMD. The χ_3 dihedral angles were obtained from Gln side chains that were not significantly solvent exposed, so as to best simulate the interior of a fibril environment. The extent of β-sheet dissociation was determined qualitatively, by viewing the trajectories and monitoring the dissociation of peptides. This qualitative analysis was also paired with a quantitative root mean square

deviation (RMSD) metric (see APPENDIX D for details).

The Gibbs free energy difference between the Figure 5.1 model *a* antiparallel and model *b* parallel β-strand fibril structures was calculated using a Python implementation of the Bennett acceptance ratio method, called Pymbar [207]. Potential energies were obtained from the NAMD log output files and used as inputs for Pymbar via a Python script that can be found in APPENDIX D.

Hydrogen bonding analysis on models *a* and *b* (from Figure 5.1) was also done by using VMD 1.9.1. Hydrogen bond contacts were defined by a heavy atom (N - - - O) distance of $<3.0\,\text{Å}$ and a N–H O angle between $-30°$ and $30°$. The number of peptide backbone-backbone, backbone-side chain, side chain-side chain, and peptide-water hydrogen bonds were calculated for models *a* and *b* over the course of the entire trajectory. The data for each hydrogen bond category was binned for each model and plotted with the statistical computing package R (see APPENDIX D).

5.4 RESULTS AND DISCUSSION

5.4.1 Q10 Forms Amyloid-like Fibril Aggregates

Figure 5.2a,b show the TEM images of NDQ10 and DQ10 aggregates. The TEM images of both NDQ10 and DQ10 aggregates resemble those of amyloid-like fibrils and exhibit morphologies similar to those of polyQ peptides with larger, more pathologically relevant Gln repeat lengths [32]. NDQ10 aggregates (Figure 5.2a) cluster into dense meshworks composed of small, thin fibrils, while DQ10 (Figure 5.2b) forms long, ribbon-like fibrils.

We used X-ray diffraction to further characterize the fibril-like nature of NDQ10 and DQ10 aggregates. Oriented amyloid fibril films exhibit a characteristic diffraction pattern with a "meridional reflection" at $\sim 4.8\,\text{Å}$ and an "equatorial reflection" at $\sim 10\,\text{Å}$ to $12\,\text{Å}$. This diffraction pattern is the hallmark of cross-β structures, wherein constituent β-strands orient perpendicular to the long axis of the fibril. The meridional reflection is indicative of the spacing between β-strands, and the equatorial reflection is indicative of the spacing

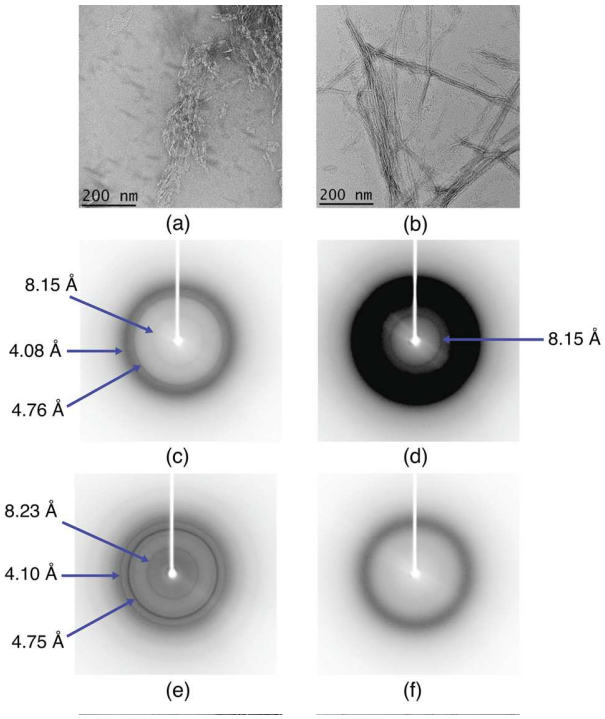


Figure 5.2: Transmission electron micrographs (TEM) and X-ray diffraction patterns of NDQ10 and DQ10 fibril aggregates. TEM of (a) NDQ10 and (b) DQ10 fibrils. The scale bars represent 200 nm. X-ray diffraction pattern of (c,d) NDQ10 fibril films, (e) DQ10 fibril films, and (f) the quartz capillary tube. The patterns shown in (c) and (d) are the same except for the contrast settings, which have been set in (d) to highlight the weak reflection at ∼8.2 Å. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

between β-sheets in the fibrils.

Figure 5.2c-e show the "powder-like" diffraction patterns of unoriented NDQ10 and DQ10 aggregate films. There are reflections at ∼4.1 Å, ∼4.8 Å, and ∼8.2 Å observed in the diffraction patterns of NDQ10 and DQ10. Similar reflections have been observed in larger sequences of polyQ fibrils in previous studies [54, 55]. Based on these previous studies, we assign the prominent ∼4.8 Å reflection to the repeat distance between neighboring β-strands within the fibril β-sheets. The weaker ∼8.2 Å reflection is assigned to the inter-β-sheet stacking repeat distance.

The equatorial reflections of most amyloid-like fibrils are very diffuse, which indicates

limited ordering and disordered growth in the inter-β-sheet dimension [208]. In contrast, polyQ fibrils usually show very sharp reflections between ~8 Å to 9 Å. We observe reflections at ~8.15 Å and ~8.23 Å in NDQ10 and DQ10, respectively. A higher order reflection occurs at ~4.1 Å for both NDQ10 and DQ10. The presence of these higher orders and the narrowness of the ~8.2 Å reflections suggests that there is long range ordering in the inter-β-sheet dimension of polyQ fibrils [54, 55]. As noted by Atkins and Sikorski [54], this long range ordering presumably arises from the tight interdigitation of the Gln side chains from neighboring sheets. In addition, there is strong inter-amide hydrogen bonding interactions between neighboring side chains of the same β-sheet.

5.4.2 UVRR of Polyglutamine Fibrils in H_2O

We utilized UVRR to investigate the molecular structure of NDQ10 and DQ10 fibrils in solution. The ~197 nm and ~204 nm excited UVRR spectra of NDQ10 and DQ10 fibrils are shown in Figure 5.3. Raman excitation at ~200 nm occurs within the NV_1 electronic transitions of secondary amide peptide bonds and to the long wavelength side on the NV_1 transitions of Gln side chain primary amide groups [125, 134]. Thus, the ~200 nm excitation UVRR spectra of polyQ peptides are dominated mainly by resonance enhanced bands that derive from primary and secondary amide (Am) vibrations, which we label with the superscripts ^{P} and ^{S}, respectively.

5.4.2.1 Assignment of Gln Side Chain Bands The main difference between the ~197 nm and ~204 nm excited UVRR spectra are the relative intensities of the primary (side chain) and secondary (peptide bond) amide bands. We previously showed that excitation at ~197 nm enhances primary amide UVRR bands significantly more than ~204 nm excitation [125]. As a result, primary amide vibrations can be highlighted by calculating the difference spectrum between the ~197 nm and ~204 nm UVRR spectra.

Figure 5.3 shows that the 197 nm – 204 nm difference spectra highlights the primary amide vibrations of the Gln side chains in NDQ10 and DQ10 fibrils. The AmI^{P} (predominately a $C_{\delta}{=}O_{\epsilon 1}$ stretching vibration) and the $AmII^{P}$ (mainly $N_{\epsilon 2}H_2$ scissoring) bands are

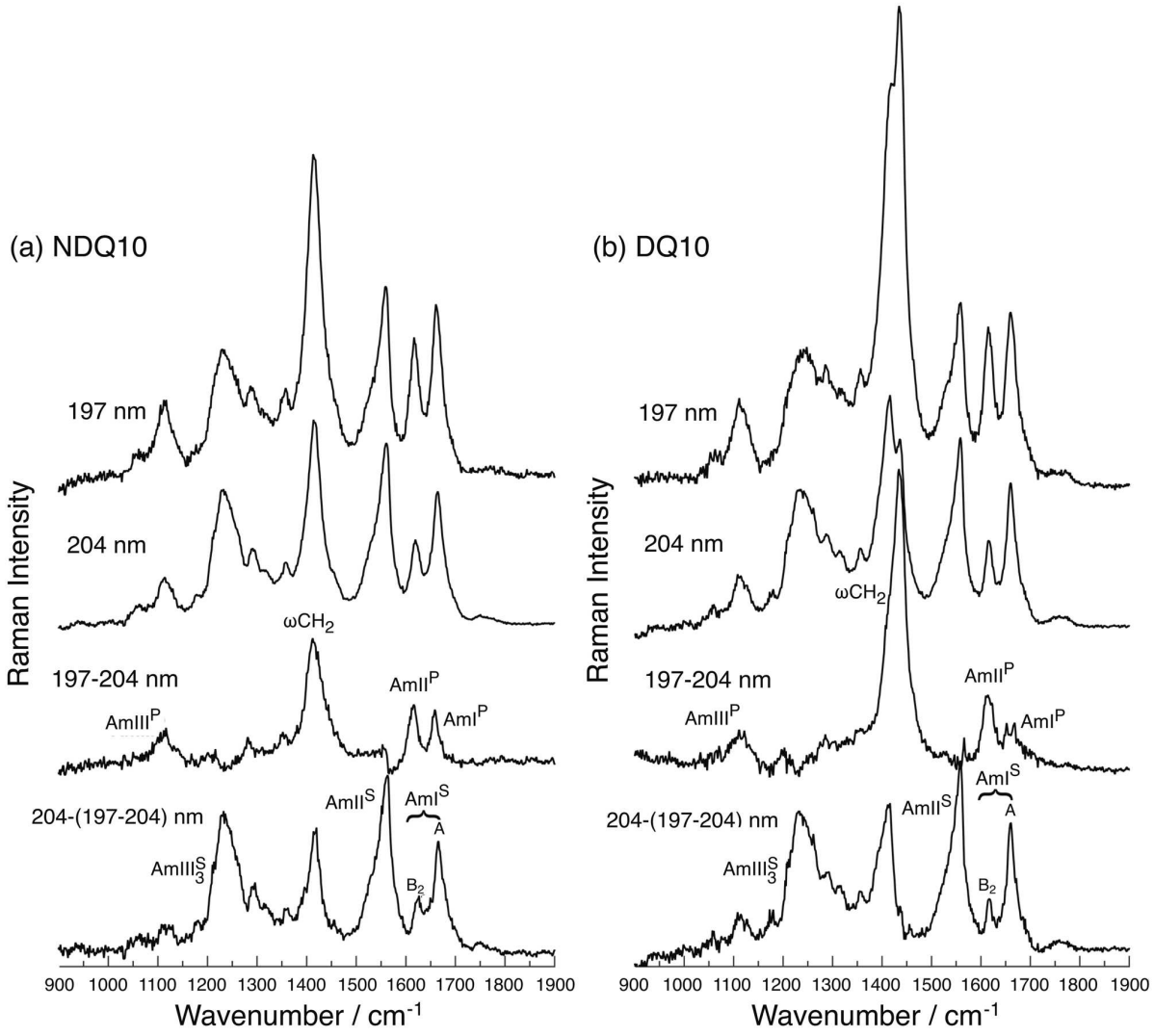


Figure 5.3: UVRR spectra (197 nm and 204 nm excitation) of (a) NDQ10 and (b) DQ10 fibrils prepared in H_2O. The spectra were measured on precipitates that were resuspended in H_2O. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

located at $\sim$1660 cm^{-1} and $\sim$1615 cm^{-1}, respectively, for both NDQ10 and DQ10 fibrils. In addition, both NDQ10 and DQ10 show a band at $\sim$1100 cm^{-1} that derives from an in-phase combination of C_δ–$N_{\epsilon 2}$ stretching and $N_{\epsilon 2}H_2$ rocking motions. As discussed in detail below, this vibration, which we call the $AmIII^P$, is sensitive to the structure of the Gln side chains.

The primary amide band frequencies and Raman cross sections are very sensitive to

the hydrogen bonding and the local dielectric environment of Gln side chains [146]. For example, the AmI^P band frequency is diagnostic of $C_\delta{=}O_{\epsilon 1}$ hydrogen bonding. The $AmII^P$ band frequency reports on hydrogen bonding of the $N_{\epsilon 2}H_2$ group. Compared to monomeric Gln in water [180], the NDQ10 and DQ10 fibril AmI^P bands are downshifted by $\sim$20 cm^{-1}, while the $AmII^P$ bands are downshifted by $\sim$7 cm^{-1}. These frequency downshifts indicate that the inter-amide hydrogen bonding within the fibrils is much stronger than the amide-H_2O hydrogen bonding that occurs for monomeric Gln. In addition, the NDQ10 and DQ10 fibril AmI^P and $AmII^P$ bands show very narrow linewidths, similar to those seen in UVRR spectra of Gln crystals [180], which indicates that the primary amide groups are in very well-defined hydrogen bonding states.

The most noticeable difference between the NDQ10 and DQ10 fibril 197 nm - 204 nm difference spectra occur for strong bands located between $\sim$1400 cm^{-1} to 1500 cm^{-1}. A band located at $\sim$1415 cm^{-1} in NDQ10 fibrils upshifts $\sim$15 cm^{-1} in DQ10 fibrils. This band derives from a complex vibration that contains CH_2 wagging, C–C stretching, CH_2 scissoring, and C_δ–$N_{\epsilon 2}$ stretching motions. Based on our previous work [146], an upshifted CH_2 wagging band signals that the Gln side chains are in an environment of higher dielectric constant. Thus, we conclude that the DQ10 Gln side chain methylene groups are in a higher dielectric constant environment than are those of NDQ10. This presumably correlates with the $\sim$0.8 Å larger inter-β-sheet spacing for DQ10 fibrils compared to NDQ10 fibrils (Figure 5.2c-e). The larger inter-sheet spacing allows solvating water molecules to penetrate deeper into DQ10 fibrils, increasing the local dielectric constant. This hypothesis is supported by the results shown in Figure D3, wherein the CH_2 wagging band downshifts to $\sim$1415 cm^{-1} and $\sim$1407 cm^{-1} in dried DQ10 and NDQ10 fibril films, respectively.

5.4.2.2 Assignment of Peptide Backbone Bands

We subtracted the 197 nm - 204 nm difference spectra from the 204 nm excited UVRR spectra in order to reveal the peptide bond secondary amide bands. The AmI^S (mainly peptide backbone C=O stretching) appears as two bands (labelled as AmI^S_A and AmI^S_{B2}) in the spectra of both NDQ10 and DQ10 fibrils. This "excitonic splitting" is diagnostic of β-sheet structures and derives from through-space transition dipole coupling between the AmI^S oscillators. The Raman spectral AmI^S splitting

patterns of antiparallel and parallel β-sheets are similar. Thus, it is usually difficult to discriminate between these two structures using this band alone [209, 210].

The intense, high frequency AmI^{S} band (labelled as the AmI$^{S}_{A}$) appears at $\sim$1665 cm^{-1} and $\sim$1660 cm^{-1} in NDQ10 and DQ10 fibrils, respectively. The less intense, low frequency AmI^{S} band (AmI$^{S}_{B2}$) appears at $\sim$1620 cm^{-1} in NDQ10 and at $\sim$1615 cm^{-1} in DQ10. These $\sim$5 cm^{-1} decreases in the AmI^{S} mode frequencies suggest slightly stronger peptide backbone C=O hydrogen bonding between β-strands in DQ10 fibrils than in NDQ10 fibrils [98]. We are, however, aware that the Raman and IR AmI^{S} bands can also be impacted by β-sheet twisting and stacking, as well as the registry of the β-strands, as described in detail by Keiderling and coworkers [209]. Thus, these AmI^{S} frequency differences between NDQ10 and DQ10 could also signal subtle differences in the twisting and stacking of the fibril β-sheets. We are continuing to examine these issues.

The extended AmIII^{S} UVRR spectral region between 1200 cm^{-1} to 1350 cm^{-1} is generally considered to be the most structurally informative [89]. This region in polyQ peptides is complicated due to the overlap of bands from Gln side chain CH_2 twisting and wagging modes that occur between $\sim$1280 cm^{-1} to 1350 cm^{-1}. In addition, the AmIII^{S} region consists of three sub-bands (called the AmIII$^{S}_{1}$, AmIII$^{S}_{2}$, and AmIII$^{S}_{3}$), which derive from vibrations that are composed of in-phase combinations of peptide bond N–H in-plane bending and C–N stretching motions.

Mikhonin *et al.* [127] previously assigned the AmIII$^{S}_{1}$, AmIII$^{S}_{2}$, and AmIII$^{S}_{3}$ bands in detail. In NDQ10 and DQ10 fibrils, the AmIII$^{S}_{1}$ occurs at $\sim$1315 cm^{-1}, while the AmIII$^{S}_{2}$ occurs at $\sim$1290 cm^{-1}, and the AmIII$^{S}_{3}$ occurs between $\sim$1200 cm^{-1} to 1280 cm^{-1}. Both NDQ10 and DQ10 show peaks at $\sim$1230 cm^{-1} and low-frequency shoulders at $\sim$1210 cm^{-1}. As discussed in detail below, the AmIII$^{S}_{3}$ band frequency is sensitive to the Ramachandran Ψ dihedral angle, and thus can be used to obtain quantitative information on the secondary structure of polyQ fibrils.

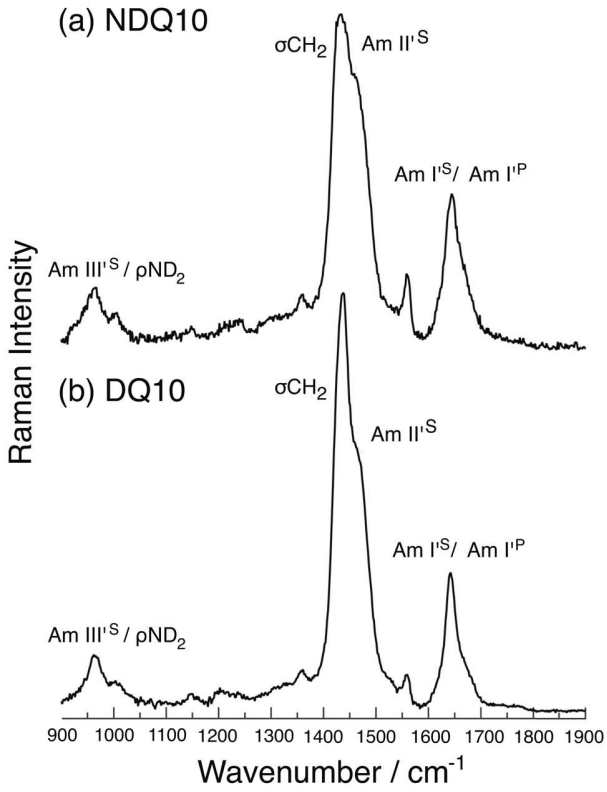


Figure 5.4: UVRR spectra (204 nm excitation) of (a) NDQ10 and (b) DQ10 fibrils prepared in D_2O. The spectra were measured on precipitates that were resuspended in D_2O. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

5.4.3 UVRR of Polyglutamine Fibrils in D_2O

Figure 5.4 shows the UVRR spectra of NDQ10 and DQ10 fibrils prepared and measured in D_2O. Deuteration of the polyQ peptide backbone N–H and primary amide side chain $N_{\epsilon 2}H_2$ groups leads to significant spectral changes. Upon N-deuteration, the AmI^S downshifts to $\sim$1640 cm^{-1} (AmI'^S) and overlaps the AmI'^P. In the case of the $AmIII^S$ mode, deuteration decouples N–H in-plane bending from C–N stretching [127]. This leaves a weak $AmIII'^S$ band (mainly N–D in-plane bending) in the $\sim$950 cm^{-1} to 1050 cm^{-1} region. The $AmIII'^S$ region also contains bands that derive from side chain $N_{\epsilon 2}D_2$ rocking modes [127]. The loss of the $AmIII^S$ band reveals the presence of several weak bands between $\sim$1300 cm^{-1} to 1400 cm^{-1}, which derive mainly from side chain CH_2 and peptide backbone C–H deformation modes.

5.4.4 Hydrogen-Deuterium (HX) Exchange of Polyglutamine Fibrils

A comparison of Figure 5.3 and Figure 5.4 show that UVRR can be employed to differentiate between N–H and N–D peptide bonds. Hydrogen-deuterium exchange (HX) can selectively probe solvent exposed versus solvent shielded peptide bonds in fibrils since the $AmIII^S$ completely disappears upon N-deuteration [91]. For example, fibrils prepared in H_2O and subjected to HX in D_2O will show $AmIII_3^S$ bands that derive mainly from peptide bonds shielded from solvent, such as those that are buried within the cross-β core. In contrast, fibrils prepared in D_2O that are subjected to HX in H_2O will show $AmIII_3^S$ bands that derive mainly from peptide bonds that are solvent accessible, generally because they are located on the aggregate surface or because they exist in exchangeable conformations such as "disordered" regions, turns, or loops.

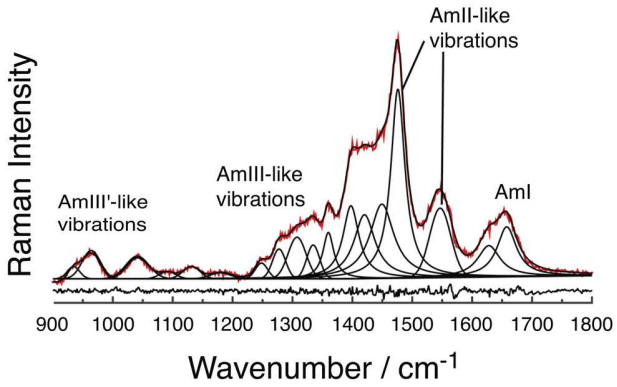


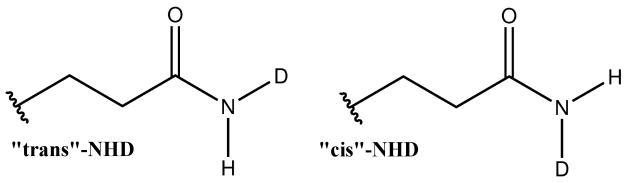
Figure 5.5: UVRR spectrum (204 nm excitation) of Gln measured in a 50%H_2O/50%D_2O mixture. The spectral contributions of solvent, as well as fully protonated and fully deuterated Gln have been subtracted. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

5.4.4.1 UVRR Bands of Partially Deuterated Primary Amides

Extensive HX of polyQ fibrils may not completely deuterate the primary amide $N_{\epsilon 2}H_2$ groups since the Gln side chains may be involved in extensive hydrogen bonding interactions. Based on Saito and coworkers' normal mode analyses of acetamide [211, 212], the partial deuteration of primary amides results in decoupling of N–H and N–D motions. As shown below, mono-deuteration

of primary amides can give rise to secondary amide-like vibrational modes! Thus, it is conceivable that mono-deuterated Gln side chains can result in AmIII^S-like vibrations.

To investigate the potential presence of AmIII^S-like vibrations in mono-deuterated primary amides, we measured the UVRR spectrum of Gln in a 50%/50% mixture of H_2O and D_2O. The spectrum, shown in Figure 5.5, was assigned with the aid of DFT calculations (see APPENDIX D for details). We considered two geometrical isomers in calculating the vibrational normal modes of partially deuterated Gln, as shown in Scheme 5.1. The "trans-NHD" Gln species resembles the trans-isomer configuration of the peptide bond, and thus, is expected to give rise to vibrations that resemble the canonical AmI^S, AmII^S, and AmIII^S vibrations of secondary amides.



Scheme 5.1: Geometric isomers of the mono-deuterated primary amide group of Gln side chains. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

We present a detailed assignment of the Figure 5.5 spectrum in APPENDIX D, as shown in Tables S1 - S3. According to our normal mode analysis, we assign an AmI^S-like vibration to an $\sim$1660 cm^{-1} band, and two AmII^S-like vibrations to bands located at $\sim$1550 cm^{-1} and $\sim$1475 cm^{-1}. The DFT calculations also indicate that the $\sim$1250 cm^{-1} to 1400 cm^{-1} region contains four vibrations with significant C_δ–$N_{\epsilon 2}$ stretching character. Two of these vibrations derive from the cis-NHD species of Gln and are assigned to the $\sim$1335 cm^{-1} and $\sim$1280 cm^{-1} bands. The other two vibrations at $\sim$1310 cm^{-1} and $\sim$1250 cm^{-1} derive from the trans-NHD species of Gln. These modes resemble AmIII^S-like vibrations since they contain a significant combination of NHD scissoring and C_δ–$N_{\epsilon 2}$ stretching motions.

5.4.4.2 HX of NDQ10 and DQ10 fibrils

Our normal mode analysis and band assignments of the Figure 5.5 Gln spectrum indicate that HX of polyQ fibrils can result in partially deuterated Gln side chains that give rise to AmIII^S-like vibrations. These AmIII^S-like bands

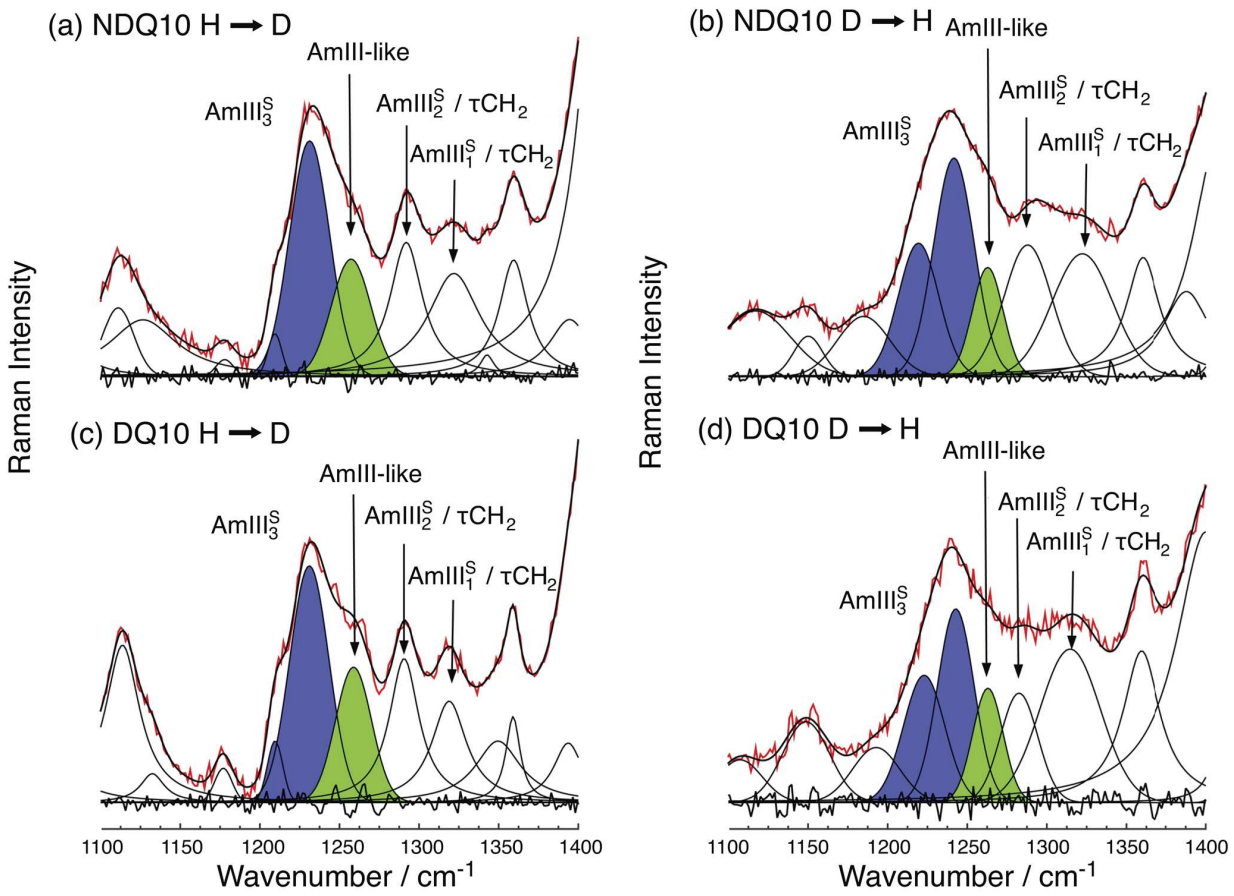


Figure 5.6: HX-UVRR spectra (204 nm excitation) of (a, b) NDQ10 and (c,d) DQ10 fibrils. For (a, c), fibrils were prepared in H_2O and exchanged in D_2O. For (b, d), fibrils were prepared in D_2O and exchanged in H_2O. The spectra were measured on precipitates that were resuspended in either H_2O or D_2O. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

appear in the high frequency side of the AmIII$_3^S$ region; however, they do not significantly overlap with that of the AmIII$_3^S$ region of β-sheets, which occurs between $\sim$1200 cm^{-1} to 1240 cm^{-1}. Thus, we can straightforwardly assign the β-sheet AmIII$_3^S$ bands of NDQ10 and DQ10 fibrils in the HX-UVRR spectra.

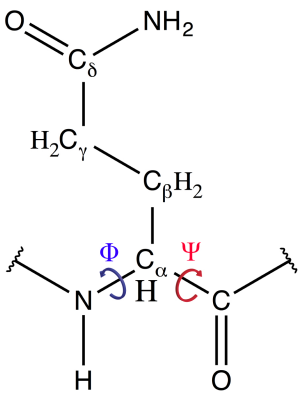
Figure 5.6 shows the curve-resolved AmIII^S region of the UVRR spectra of NDQ10 and DQ10 fibrils following HX. The spectra labeled as H→D (D→H) were measured from fibrils prepared in H_2O (D_2O) and subjected to HX in D_2O (H_2O). The bands shown in blue are

assigned to true $AmIII_3^S$ vibrations, while those shown in green are assigned to the $AmIII^S$-like vibrations that derive from partially deuterated primary amides. The H→D spectra of NDQ10 and DQ10 fibrils show $AmIII_3^S$ bands at ∼1210 cm^{-1} and ∼1230 cm^{-1}. These bands upshift ∼10 cm^{-1} in the D→H spectra. As discussed in detail by Mikhonin *et al.* [97], this peptide bond frequency upshift can result from an increased hydrogen bonding to water which does not involve any structural changes. Thus, we conclude that the Q10 fibril exterior peptide bonds are more extensively hydrogen bonded to water than those in the fibril interior.

5.4.5 Ramachandran Ψ Angle Distributions

The $AmIII_3^S$ band is the most conformationally sensitive secondary amide band because its frequency depends sinusoidally on the Ramachandran Ψ dihedral angle (see Scheme 5.2) of the peptide backbone [96, 97]. This sinusoidal dependence derives from the coupling of the C_α–H bending vibration with the N–H bending component of the $AmIII_3^S$ vibration. For example, this coupling is strong for β-strand and PPII-like peptide bond Ψ angles, when the C_α–H and N–H groups are in an approximately cis-configuration. The strong coupling between N–H and C_α–H bending motions downshifts the $AmIII_3^S$ frequency. In contrast, for α-helical-like Ψ angles, the C_α–H and N–H groups are in a trans-configuration to each other, which results in the C_α–H and N–H bending motions decoupling. This decoupling results in the $AmIII_3^S$ band upshifting.

We utilized the structural sensitivity of the $AmIII_3^S$ band to determine Ramachandran Ψ dihedral angle distributions for the NDQ10 and DQ10 fibril peptide bonds (Figure 5.7). To do this, we employed the methodology of Asher and coworkers [97, 105] (see APPENDIX D for details), which correlates the different frequencies of the $AmIII_3^S$ band envelope to different peptide bond Ψ angles. This enables us to determine a probability distribution of peptide bond Ψ angles from the inhomogeneously broadened $AmIII_3^S$ bandshapes shown in Figure 5.6. The Ψ distributions shown in black derive from the H→D HX-UVRR (Figure 5.6a,c) $AmIII_3^S$ band profiles, while those shown in blue are from the D→H HX-UVRR (Figure 5.6b, d) $AmIII_3^S$ band profiles.



Scheme 5.2: Depiction of Ramachandran Ψ and Φ dihedral angles in polyQ peptides. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

As shown in Figure 5.7, the Ψ angle distributions of NDQ10 and DQ10 are similar, which indicates that the fibril secondary structures are essentially the same. All the distributions are bimodal, showing peaks near ∼145° and ∼125°. According to Hovmöller *et al.*'s analysis [213] of protein data bank crystal structures, Gln residues in antiparallel β-sheet conformations show an average Ψ angle of ∼137(15)°, while those in parallel β-sheet structures show an average Ψ angle of ∼129(15)°. Given these values, we attribute the NDQ10 and DQ10 peaks centered at ∼145° to fibril peptide bonds that are in antiparallel β-sheet conformations, and the peaks at ∼125° to peptide bonds in parallel β-sheet conformations.

A comparison of the black and blue Ψ angle distributions indicates that there are structural differences between solvent accessible and inaccessible peptide bonds in both NDQ10 and DQ10 fibrils. The blue distributions are much broader than the black distributions, which indicates that the solvent accessible peptide bonds exhibit greater conformational heterogeneity than the solvent inaccessible peptide bonds. Most of this increased structural heterogeneity stems from the peptide bonds in parallel β-sheet conformations. Indeed, the standard deviations (σ) of the antiparallel β-sheet distributions corresponding to solvent accessible and inaccessible peptide bonds are not significantly different. However, the standard deviations of the parallel β-sheet distributions are ∼12°-13° for solvent accessible peptide bonds, but collapse to a narrower range of Ψ angles for solvent inaccessible peptide bonds.

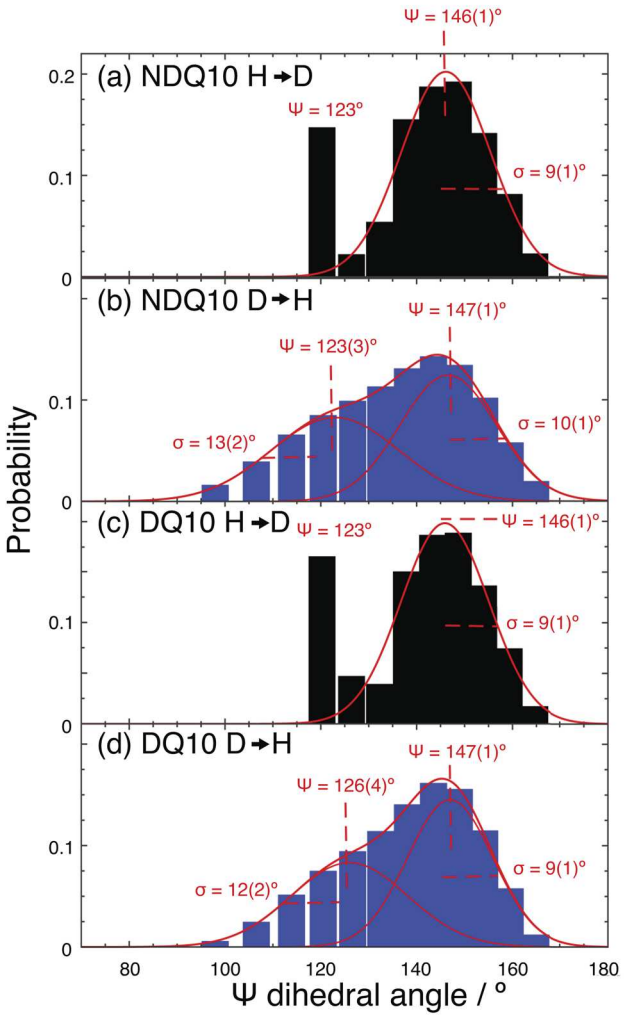


Figure 5.7: UVRR determined Ψ angle distributions for NDQ10 and DQ10 fibrils. (a) Distribution corresponding to the NDQ10 fibril Figure 5.6a HX spectrum. (b) Distribution corresponding to NDQ10 fibril Figure 5.6b HX spectrum. (c) Distribution corresponding to DQ10 fibril Figure 5.6c HX spectrum. (d) Distribution corresponding to DQ10 fibril Figure 5.6d HX spectrum. The distributions were least-squares fit to one or two Gaussians. The peak Ψ angles and distribution standard deviations (σ) are reported, along with their standard errors from the fits (in parentheses). Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

The Ψ angle distributions shown in Figure 5.7 indicate that the solvent inaccessible peptide bonds in NDQ10 and DQ10 fibrils preferentially adopt antiparallel over parallel

β-sheet conformations. The solvent inaccessible peptide bonds derive primarily from the fibril interiors, where primary fibril nucleation occurs. This suggests that nascent polyQ (proto)fibrils form around antiparallel β-sheet nuclei. In contrast, peptide bonds that are solvent accessible are located predominately on the surface of polyQ aggregates, which is more disordered since peptides can aggregate onto the fibril by forming parallel β-sheet structures, in addition to adding as antiparallel β-sheets.

5.4.6 Developing a Molecular-level Structural Model

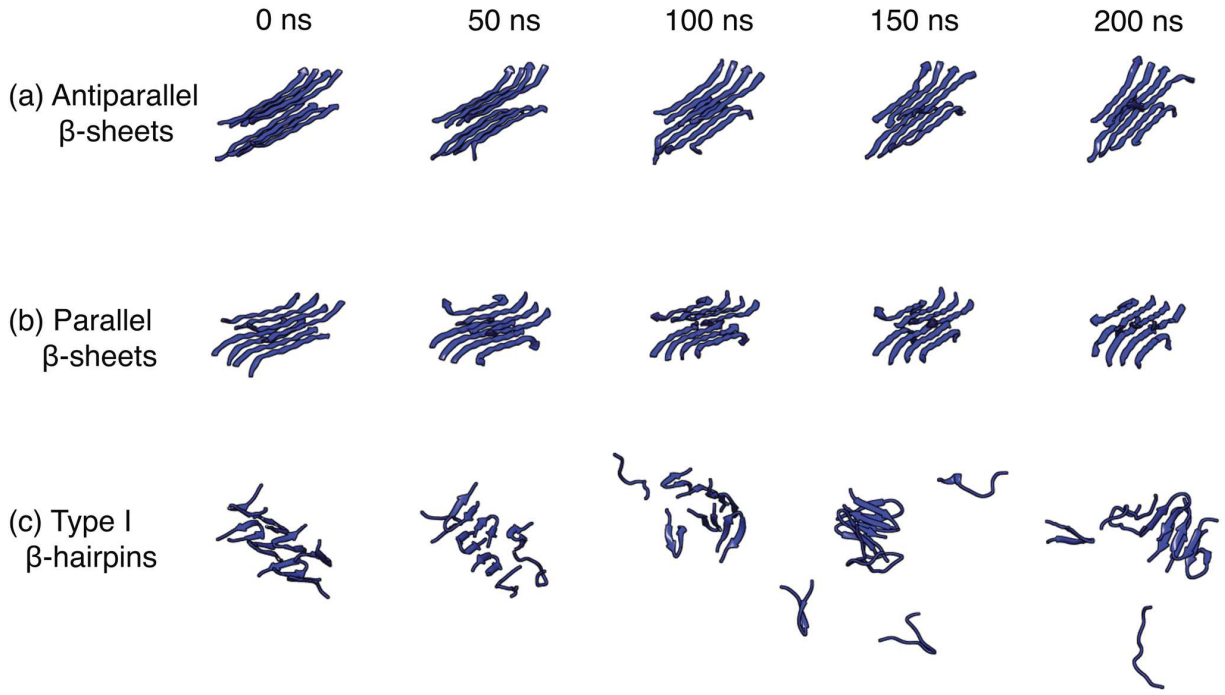


Figure 5.8: Time evolution of (a) antiparallel β-sheet, (b) parallel β-sheet, and (c) β-hairpin fibril models in MD simulations. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

The Figure 5.7 distributions show that no peptide bonds populate β-turn Ψ angles. Thus, we conclude that the basic structural motifs of NDQ10 and DQ10 fibrils are extended β-strands. To investigate this hypothesis, we utilized atomistic MD to examine Q10 peptides arranged in three different fibril architectures. As shown in Figure 5.1, these models are composed of eight Q10 peptides arranged into two β-sheet layers. Models *a* and *b* consist of extended β-strands that are arranged into antiparallel and parallel β-sheet architectures,

respectively. Model *c* consists of β-hairpins with Type I turn structures that are arranged into an antiparallel β-sheet configuration. Further details regarding the construction of the fibril models for the MD simulations can be found in APPENDIX D.

The Figure 5.1 simulated fibril models retained structural integrity throughout the energy minimization and equilibration processes of the computations. We utilized an RMSD metric (Figure D2) to monitor the extent of dissociation of the fibril models during the simulation production runs. The RMSD of peptide backbone atoms relative to that of the respective initial, energy-minimized structure was used. An RMSD of 3 Å was used as a dissociation threshold. Based on this criterion, the Figure 5.1c β-hairpin model dissociates at ∼60 ns into the production simulation, while models *a* and *b* remain intact throughout the full 200 ns simulation (Figure 5.8).

As shown in Figure 5.9, we compared the Ramachandran Ψ angle distributions obtained from the production runs of the simulated fibril models to those measured experimentally. The distributions corresponding to models *a* and *b* show large peaks at ∼141° and ∼127°, respectively, which are very close to the experimentally observed Ψ angle peaks for the antiparallel and parallel β-sheets. In contrast, the model *c* distribution shows very poor agreement with the experimentally determined distributions since the calculated peak Ψ angle distribution is downshifted ∼12° from the experimentally observed antiparallel β-sheet peak distribution. In addition, the model *c* Ψ angle distribution shows a doublet located at ∼−19° and ∼−43° that is not experimentally observed. This doublet corresponds to Ψ angles that derive from the i+1 and i+2 Type I β-turn residues. The strong agreement between the model *a* and *b* Ψ angle distributions with those measured experimentally supports our conclusion that the basic structural motif of NDQ10 and DQ10 fibrils are stacked β-strands organized into β-sheets.

Our β-strand models for NDQ10 and DQ10 fibrils are consistent with other studies. For example, Schneider *et al.* [56] suggest, on the basis of EM and solid-state NMR, that $D_2Q_{15}K_2$ peptides adopt extended β-strands in fibrils. In another study, Thakur and Wetzel [69] probed polyQ fibril structure by replacing Gln-Gln residue pairs with Pro-Gly pairs to increase the formation of β-turn structures. They found that peptides, which had continuous stretches of ∼9 to 10 Gln residues, mimicked the aggregation behavior and morphologies of

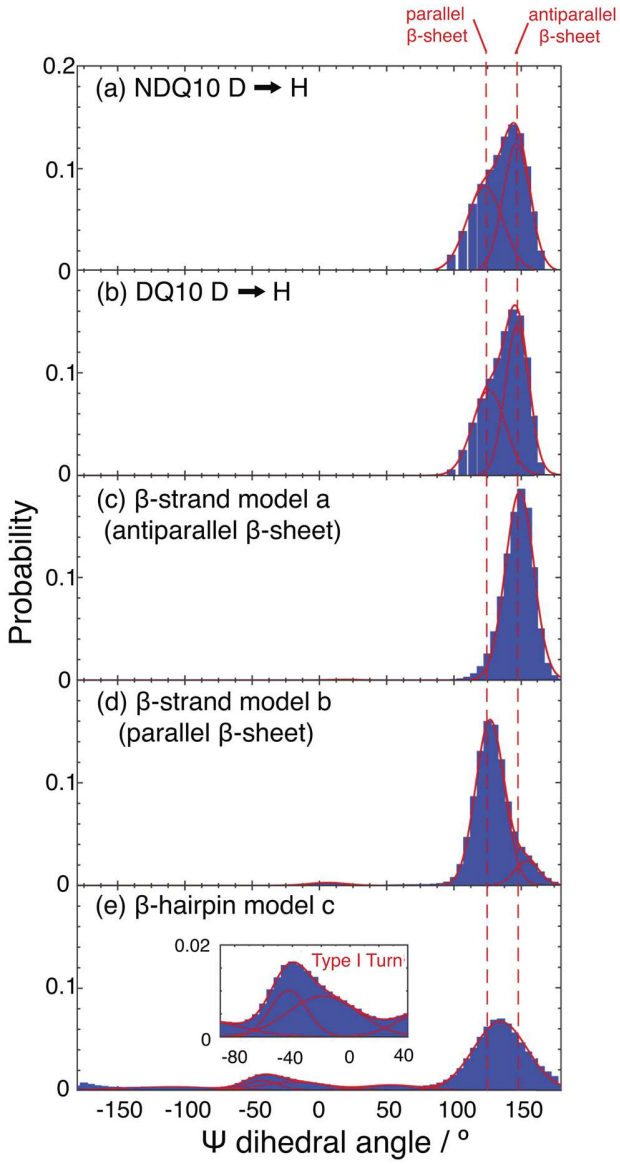


Figure 5.9: Comparison of Ψ angle distributions from simulations and experiments. The distributions correspond to: (a, b) solvent accessible peptide bonds from NDQ10 (Figure 5.7b) and DQ10 fibrils (Figure 5.7d), respectively; (c) model *a* in Figure 5.1; (d) model *b* in Figure 5.1; (e) model *c* in Figure 5.1. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

unmutated fibrils. They interpret these results to mean that simple polyQ peptides form β-strands that are optimally ~9 to 10 Gln residues in length.

5.4.7 Antiparallel β-sheets are more Favorable than Parallel β-sheets in polyQ fibrils

Numerous studies have shown that the most common structural motif of typical prions and amyloid fibrils is the in-register parallel cross-β structure [81, 82, 167, 168, 214]. For these fibrils, in-register parallel β-sheets appear to maximize hydrophobic and steric zipper interactions [198]. Reports of antiparallel β-sheets structures are far less common, although they have been reported for some amyloidogenic peptide microcrystals [215] and fibrils [83]. PolyQ fibrils are unique in that most studies [55, 56, 58, 69] suggest that they predominately form antiparallel β-sheets.

This preference indicates that antiparallel β-sheet structures is be more energetically favorable than parallel β-sheets in polyQ fibrils. To examine this possibility, we used the Bennett acceptance ratio method (see APPENDIX D, Figure D4) to calculate the free energy difference between our simulated parallel and antiparallel β-sheet fibrils. The free energy of our simulated antiparallel β-sheet model system (model *a*) was found to be $160.5(20)\,\mathrm{kJ\,mol^{-1}}$ (*per fibril system*) lower than the parallel β-sheet model system (model *b*). Since the simulated fibril systems were composed of eight Q10 peptides, this means that the free energy is $\sim$1.5 $\mathrm{kJ\,mol^{-1}}$ *per peptide bond* (there are 104 total peptide bonds in each simulated fibril system) lower for the antiparallel β-sheet structure than the parallel β-sheet structure.

The MD simulation results suggest that antiparallel β-sheet formation is favored in polyQ fibrils. This accounts for the greater fraction of antiparallel β-sheets over parallel β-sheets experimentally observed in NDQ10 and DQ10 fibrils. Using the integrated areas of the AmIII_3^S bands in the Figure 5.6 spectra, we calculate the apparent Gibbs free energy difference between parallel and antiparallel β-sheets to be $\sim$6-7 $\mathrm{kJ\,mol^{-1}}$ per solvent inaccessible peptide bond and $\sim$1 $\mathrm{kJ\,mol^{-1}}$ per solvent accessible peptide bond. We note that the experimentally measured free energy difference of 1 $\mathrm{kJ\,mol^{-1}}$ per solvent accessible peptide bond is very close to the value of $\sim$1.5 $\mathrm{kJ\,mol^{-1}}$ per peptide bond calculated from the simulated, well-hydrated fibril models.

We examined electrostatic and hydrogen bonding interactions of models *a* and *b* in order to understand the origins of the energetic favorability of antiparallel β-sheets over parallel

β-sheets. In the case of electrostatics, we find that models a and b are both stabilized by favorable electrostatic interactions between N-terminal Asp residues and C-terminal Lys residues from opposing β-sheets. However, within a β-sheet, the antiparallel β-sheet model a is stabilized by favorable inter-strand electrostatic interactions between oppositely charged terminal residues. In contrast, the parallel β-sheet model b is slightly destabilized due to inter-strand electrostatic repulsions of similarly charged terminal residues. The inter-strand repulsion between like-charged terminal residues in model b may disrupt nearby peptide backbone hydrogen bonding interactions.

Most of the energetic favorability of the antiparallel β-sheet derives from hydrogen bonding interactions. In our analysis, we considered three different types of peptide-peptide hydrogen bonding: peptide backbone-backbone, side chain-backbone, or side chain-side chain (see Figure D5 and Table D4 in APPENDIX D). Our analysis indicates that, on average, model a forms more peptide-peptide hydrogen bonds than does model b. Specifically, model a forms significantly more hydrogen bonds between peptide backbone amides than does model b (Figure D5d). In contrast, model b forms more side chain-backbone and peptide-water hydrogen bonds than does model a, as shown in Figure D5b,c. Thus, it appears that model a is stabilized by more peptide-peptide hydrogen bonds, and less destabilized due to fewer peptide-water hydrogen bonds. In contrast, model b is less energetically favorable due to fewer stabilizing peptide-peptide hydrogen bonds and more destabilizing peptide-water hydrogen bonds. We also note that it is well-known that antiparallel β-sheets are enthalpically more favorable than are parallel β-sheets structures [216] due to their optimal hydrogen bonding geometries. Our MD results support the hypothesis that thermodynamics, not kinetics, drive polyQ aggregation into antiparallel β-sheet architectures.

5.4.8 Structure of NDQ10 and DQ10 Gln side chains

5.4.8.1 Determination of Gln Side Chain χ_3 Angle Distributions We recently discovered [180] a vibrational spectral marker band that we call the AmIII^P band, which shows a cosinusoidal frequency dependence on the O–C–C–C dihedral angles of Gln and Asn side chains (the χ_3 and χ_2 angles, respectively). The AmIII^P vibration is somewhat

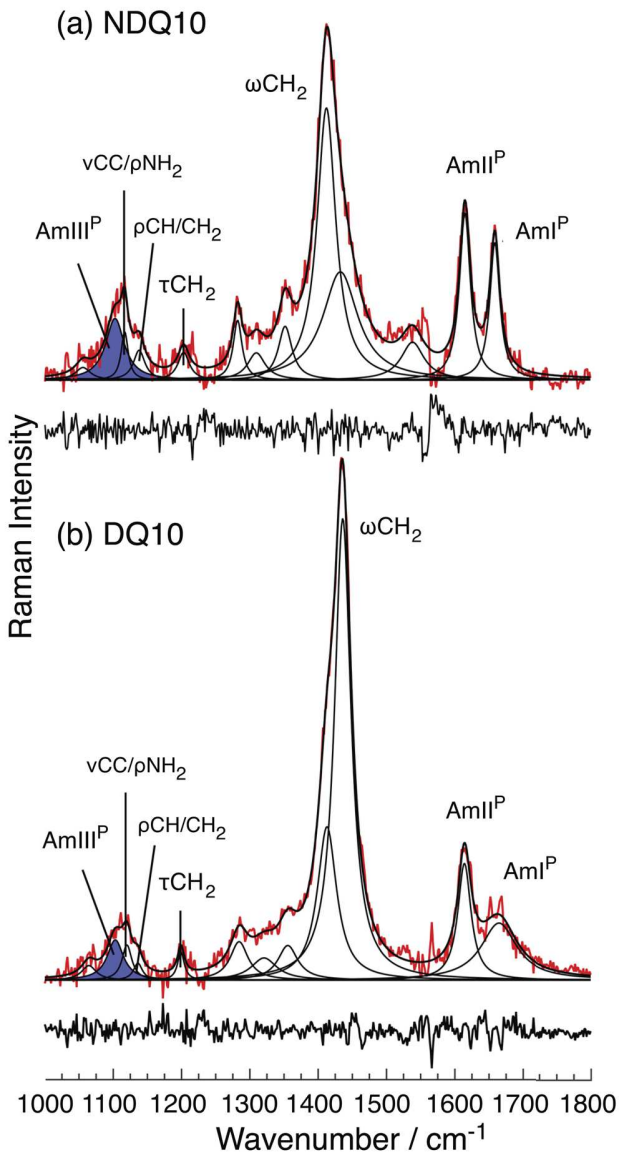


Figure 5.10: Peak fitting of the 197 nm – 204 nm UVRR difference spectra of (a) NDQ10 and (b) DQ10 fibrils. Also shown are the residuals between the fitted and measured spectra. The bands shown in blue are assigned to the $\mathrm{AmIII^P}$ of the Gln side chains. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

reminiscent of the $\mathrm{AmIII_3^S}$ vibration since it derives from an in-phase combination of $\mathrm{C_\delta}$–$\mathrm{N_{\epsilon 2}}$ stretching and $\mathrm{N_{\epsilon 2}H_2}$ rocking motions (replacing N–H bending in the $\mathrm{AmIII_3^S}$). However, $\mathrm{C_\beta}$–$\mathrm{C_\gamma}$ stretching also contributes significantly to the $\mathrm{AmIII^P}$ vibrational potential energy distribution.

The structural sensitivity of the AmIII^{P} band derives mainly from the hyperconjugation of the C_{δ}=$O_{\epsilon 1}$ π^* and C_{β}–C_{γ} σ orbitals [180]. When hyperconjugation is strong (e.g. at $\chi_3 \sim \pm 90°$) electron density is transferred from the C_{β}–C_{γ} to the C_{δ}=$O_{\epsilon 1}$ bond. This elongates the C_{β}–C_{γ} bond and reduces the C_{β}–C_{γ} stretching force constant. As a result, the AmIII^{P} frequency downshifts. In contrast, in the absence of hyperconjugation (e.g. at $\chi_3 \sim 0°$), the C_{β}–C_{γ} bond length shortens and the AmIII^{P} frequency upshifts.

To locate the AmIII^{P} bands of NDQ10 and DQ10 fibrils, we curve-resolved the 197 nm - 204 nm difference spectra. As shown in Figure 5.10, the AmIII^{P} band is located between $\sim$1050 cm^{-1} and 1150 cm^{-1}. Curve-resolving this spectral region for both NDQ10 and DQ10 reveals four underlying bands located at $\sim$1060 cm^{-1}, $\sim$1100 cm^{-1}, $\sim$1120 cm^{-1}, and $\sim$1140 cm^{-1}. Based on our previous work [180], we assign the 1060 cm^{-1}, 1120 cm^{-1}, and 1140 cm^{-1} bands to C–N stretching, C_{β}–C_{γ} stretching/$N_{\epsilon 2}H_2$ rocking, and CH_2 twisting vibrations, respectively. The AmIII^{P} band is assigned to the $\sim$1100 cm^{-1} shoulder feature [180].

We utilized the structural sensitivity of the AmIII^{P} vibration to determine the distributions of χ_3 dihedral angles for the NDQ10 and DQ10 fibrils. The methodology employed to calculate the χ_3 angle distributions is similar to that used to determine the peptide bond Ramachandran Ψ angle distributions (see APPENDIX D on for details). As shown in Figure 5.11a and b, the χ_3 distributions are doubly peaked since the AmIII^{P} band frequencies gives rise to two physically possible χ_3 angle solutions (see the discussion of eq. D.8 in APPENDIX D). The AmIII^{P} band of NDQ10 gives rise to distributions of χ_3 angles centered at $\chi_3 \sim -14°$ or 5°. The AmIII^{P} band of DQ10 gives rise to similar χ_3 angle distributions that are centered at $\chi_3 \sim -12°$ or $\sim$3°. The Gaussian-like distributions of both NDQ10 and DQ10 show standard deviations of $\sim$15°.

We compared the Figure 5.11a,b distributions to those calculated from the MD simulated fibril models. The χ_3 dihedral angle distributions corresponding to the simulated fibril models are shown in Figure 5.11c-e. The antiparallel β-sheet model *a* shows a dominating peak at $\sim$4°, whereas the parallel β-sheet model *b* shows its largest peak centered at $\sim -10°$. These dihedral angle maxima of models *a* and *b* are very close to the experimentally measured for NDQ10 and DQ10.

Our combined UVRR and validated MD results provide new detailed insights into the structure of the Gln side chains in polyQ fibrils. The experimentally measured χ_3 dihedral angles for both NDQ10 and DQ10 are distributed around $\sim$0°. As shown by the structures in Figure 5.12a,b, the model *a* and *b* Gln side chains are approximately planar. This allows Gln side chains from opposing β-sheets to form tightly interdigitated steric zippers, as well as to enable the primary amide groups to both accept and donate hydrogen bonds between neighboring β-strands.

5.4.8.2 Comparisons of Side Chain Structures with Other Models To our knowledge, the structure of Gln side chains in polyQ fibrils have only been investigated previously by the Sikorski [54] and Kirschner [55] groups. Both of these studies examined polyQ fibril structure with X-ray diffraction. However, despite obtaining similar diffraction patterns, the Sikorski [54] and Kirschner [55] groups proposed different fibril structures based upon modeling the data. For example, the Sikorski [54] group proposes that Gln side chains show alternating χ_1 (N–C_α–C_β–C_γ) torsion angles along each β-strand of $\sim$69° and $\sim-$113°. The Sikorski group's model [54] is similar to our structure since the Gln side chains are approximately planar; however, their final model indicates that the χ_3 angle is $\sim$180°, which is energetically unfavorable and sterically nearly forbidden [180].

The Kirschner [55] group proposed a different structure for Gln side chains in polyQ fibrils. In their study, an electron density maps were calculated from their low-resolution powder-like X-ray diffraction patterns by combining the structure factors determined from the experimental reflection intensities, and the phase angles from a model that satisfactorily accounted for the observed d-spacings. They then modeled the electron densities to generate atomic models for $Ac{-}Q_8{-}NH_2$ and $D_2Q_{45}K_2$ fibrils. Their modeling of the electron density maps suggests that polyQ fibril Gln side chains adopt unusual bent conformations, as shown in Figure 5.12c,d. It should be noted that the fibril models reported by the Kirschner [55] group result in very high R-factors (24% for $Ac{-}Q_8{-}NH_2$ and 35% for $D_2Q_{45}K_2$).

The Kirschner [55] group's side chain structures are inconsistent with our model. The Kirschner [55] group structure shows side chain χ_3 angles that approach values of $\sim \pm$90°, which differs significantly from our experimentally determined values that are close to $\sim$0°.

The fact that we measure χ_3 angles near $\sim 0^\circ$ is important because it means that the Gln side chains are roughly planar, which allows the steric zipper interactions that are believed to stabilize amyloid-like fibrils. It is difficult to envision Gln primary amide stabilizing steric zipper interactions in the Kirschner [55] group model, since their Gln side chains conformations are bent, and should not enable tight interdigitation of neighboring β-sheets.

5.4.9 Dependence of Fibril Structure on Deposition of Different NDQ10 and DQ10 Solution Conformations

The NDQ10 and DQ10 peptides have significantly different solution conformations. The NDQ10 and DQ10 peptide solutions are composed of putative β-hairpin-like and PPII-like conformations, respectively [125]. These two solution conformations aggregate into fibrils that are composed of similar, but not identical, β-sheet structures, which have similar planar Gln side chain conformations. The similarity between these fibril structures probably signals that the extended β-strand structure is the most energetically favorable Q10 fibril structure.

However, the subtle differences observed between the NDQ10 and DQ10 fibril structures must result from the variation in the growth mechanisms due to different Q10 solution conformations. A speculative hypothesis is that the increased water content and spacing of the DQ10 fibrils results from the preferential addition of the well hydrated PPII-like DQ10 peptides that lead to incorporation of water into the hydrophobic interdigitating side chain domains. This gives rise to the longer DQ10 fibril inter-β-sheet spacing compared to NDQ10, as observed in Figure 5.2. We are presently investigating this possibility.

5.5 CONCLUSION

We performed a detailed structural analysis of NDQ10 and DQ10 fibrils using UVRR and MD simulations. On the basis of comparing our UVRR and MD simulation results, we determine that the basic structural element of Q10 fibrils is an extended β-strand. The solvent inaccessible interiors of NDQ10 and DQ10 fibrils are a predominately antiparallel β-sheet

structures that are highly ordered and composed of these extended β-strands. However, the water accessible peptide bonds, which are located predominately on the fibril surfaces, show greater conformational heterogeneity and contain significant subpopulations of β-strands that adopt parallel β-sheet architectures.

Our MD simulation results indicate that Q10 antiparallel β-sheets are energetically more favorable than parallel β-sheets. This is an important insight because it may explain why polyQ fibrils, in contrast to most amyloid-like aggregates, preferentially adopt antiparallel β-sheets instead of in-register parallel β-sheets. Our results indicate that the origin of the energetic favorability of Q10 fibril antiparallel β-sheets is mainly due to hydrogen bonding. Antiparallel β-sheets form, on average, more hydrogen bonds between peptide backbone amides than do parallel β-sheets.

This study also provides important new insights into the structure and chemical environment of Gln side chains in polyQ fibrils. In contrast to previous, low-resolution X-ray studies, we show that that the Gln side chains in polyQ fibrils adopt conformations that are roughly planar, where the χ_3 dihedral angles are narrowly distributed around 0°. This enables the formation of steric zippers, wherein the side chains of neighboring β-sheets tightly interdigitate.

The UVRR spectra are also highly sensitive to the local hydrogen bonding and dielectric environments of the Gln side chains. For example, a major difference observed between NDQ10 and DQ10 fibrils is a different local dielectric environment of the Gln side chains. The primary amides of both NDQ10 and DQ10 fibrils are strongly hydrogen bonded; however, in DQ10, the side chain methylene groups experience a higher dielectric constant environment. This is likely correlated with the larger DQ10 fibril inter-sheet spacing compared to NDQ10. The larger inter-sheet spacing of DQ10 fibrils presumably results from an increased content of water between β-sheets. The NDQ10 and DQ10 fibrils experience different growth processes due to their different Q10 solution conformations. Deposition of the well-hydrated PPII-like DQ10 peptides results in fibrils with higher water content, with a β-sheet structure showing larger inter-sheet spacings than occurs for growth with β-hairpin NDQ10 solution peptides.

Finally, our study demonstrates the utility of synergistically coupling UVRR with MD simulations. Understanding the structure of polyQ and other amyloid-like fibrils remains

of great importance because these aggregates are implicated in numerous neurodegenerative diseases. Knowing the fibril structures will provide important insights into the aggregation mechanism(s) of polyQ peptide sequences. However, determining the molecular-level structure of fibrils is challenging because of the insoluble and non-crystalline nature of fibril aggregates. A key advantage of our approach is that UVRR can be utilized to measure the peptide backbone Ψ and the side chain χ_3 dihedral angle distributions. This information can be used to generate hypotheses on the structure of fibrils, which can then be tested *in silico* with MD simulations. Our work combines UVRR and MD into a novel approach for investigating fibrils. Our approach complements existing methods such as solid-state NMR. However, the short experimental UVRR acquisition time frames of our measurements enable the examination of peptide structural changes during fibril aggregation on timescales that are inaccessible to conventional biophysical methods.

5.6 NOTE ADDED IN REVISION

While this paper was in review, Hoop et al. [61] published a study that examined fibrils prepared from the huntingtin exon 1 domain (htt exon 1). Overall, their solid-state NMR (ssNMR) measurements on the polyQ fibril core of htt exon 1 are in remarkable agreement with our UVRR-based measurements on Q10. Their findings suggest that the polyQ fibril core in htt exon 1 is arranged in β-hairpins that form antiparallel β-sheets. Using sophisticated magic angle spinning ssNMR techniques, Hoop et al. measure Ramachandran Ψ angles that are very close to our values for antiparallel β-sheets. This is particularly true for their observed "b-type" conformer, where they measure Ψ angles of $\sim$150°.

Hoop et al. also measured the Gln side chain χ_2 dihedral angles (C_α–C_β–C_γ–C_δ) to be $\sim$180° in htt exon1 fibrils. Their results lead them to also conclude that the Gln side chains in polyQ fibrils are extended in structure. Although they did not directly measure Gln side chain χ_3 angles, Hoop et al. suggest values of $\pm$150°, which differ significantly from our *experimentally determined* values reported here. It is interesting to note that our experimentally validated MD-simulated β-sheet fibril structure (model *a*) shows a mean χ_2

value of ∼180°, which is exactly the angle that Hoop et al. measure. Combining these ssNMR results with our UVRR and MD data, leads us to propose that the most likely fibril structures of the Gln side chains will have χ_2 and χ_3 angles of ∼180° and ∼0°, respectively. We are examining this issue in greater detail.

5.7 ACKNOWLEDGMENTS

Funding for this work was provided by the University of Pittsburgh (D.P., Z.H., S.A.A.) and partially supported by NIH R01 DA027806 (J.D.M. and R.J.W.). The MD simulation computer time was supported by XSEDE MCB060069, and computer equipment was purchased from NSF funds (CHE-1126465 and P116Z080180). The DFT calculations were supported by the University of Pittsburgh Center for Simulation and Modeling through the super-computing resources provided. We thank Dr. Sergei V. Bykov, Liqi Feng, Jonathan Weisberg, and Jonathan Wert for useful discussions. We are also grateful to Dr. Steven Geib and Dr. Alexander Makhov for technical assistance with the X-ray diffraction measurements and electron microscopy, respectively.

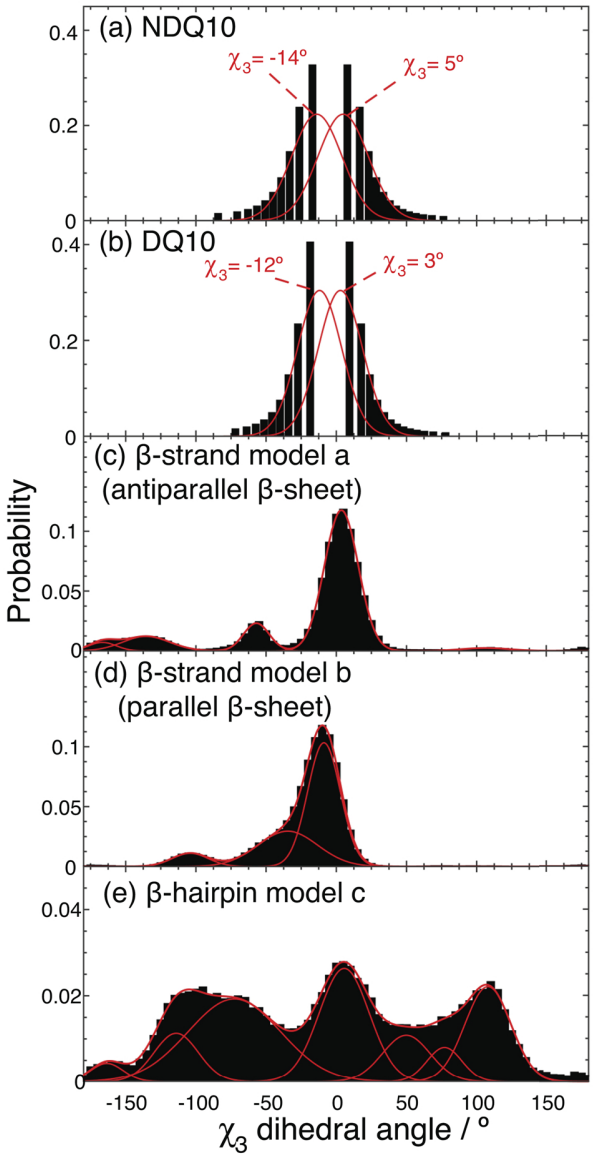


Figure 5.11: χ_3 dihedral angle distributions of Gln side chains. Distributions obtained experimentally are shown for (a) NDQ10 fibrils and (b) DQ10 fibrils. Distributions obtained from MD simulated structures correspond to (c) model *a* in Figure 5.1; (d) model *b* in Figure 5.1; and (e) model *c* in Figure 5.1. The missing χ_3 angles in (a,b) around the region of $\sim 0^\circ$ is due to the fact that those corresponding $\mathrm{AmIII^P}$ frequencies are outside the domain of the semi-empirically derived equation [180] used to correlate the frequencies of the $\mathrm{AmIII^P}$ band envelope to their respective χ_3 angles (see eq. D.8 in APPENDIX D). Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

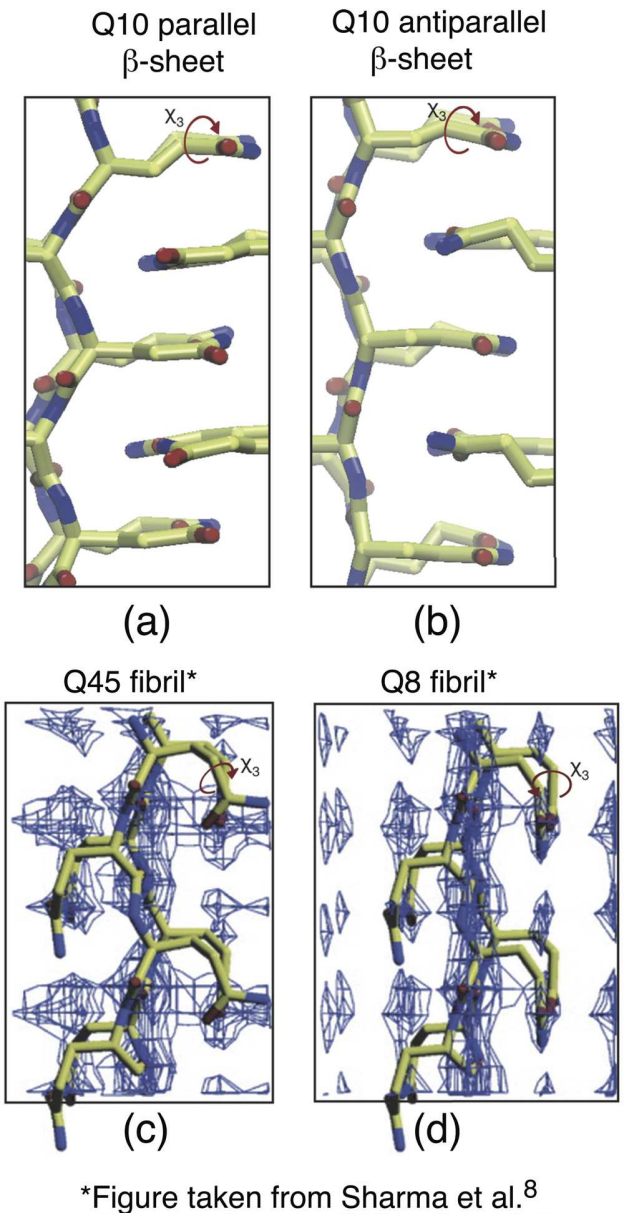


Figure 5.12: Comparison of side chain geometries from (a,b) our MD simulations and (c,d) Sharma *et al.* [55] The (c,d) figure panels were adapted from reference [55]. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

APPENDIX A

DERIVATIONS

A.1 DERIVATION OF THE TRANSITION SCATTERING RATE

A.1.1 Second Order Perturbation Theory

In this section, the transition scattering rate for Rayleigh and Raman (eq. 2.1) is derived using second order perturbation theory in the interaction picture of quantum mechanics. We begin by expanding the Dyson series to its second order term [113]:

$$\begin{aligned}\hat{U}_I(t,0) &= 1 + \left(\frac{1}{i\hbar}\right)\int_0^t dt_1 \hat{H}_I(t_1) + \left(\frac{1}{i\hbar}\right)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \hat{H}_I(t_1)\hat{H}_I(t_2) \\ &= \hat{U}_I^{(0)}(t,0) + \hat{U}_I^{(1)}(t,0) + \hat{U}_I^{(2)}(t,0)\end{aligned} \tag{A.1}$$

where $\hat{U}_I(t,0)$ is the time evolution operator. The operator, $\hat{H}_I(t)$, is the interaction picture Hamiltonian [113], which is defined as:

$$\hat{H}_I(t) = e^{i\hat{H}_0 t/\hbar}\hat{H}'(t)e^{-i\hat{H}_0 t/\hbar} \tag{A.2}$$

$\hat{H}_0$ is the time-independent Hamiltonian of the system. $\hat{H}'(t)$ is the time-dependent Hamiltonian due to the interaction between the molecule being excited and the perturbing radiation field. The perturbation is small and acts over the time interval from $t = 0$ to t. For simplicity, this perturbation is treated classically. We assume that the perturbation is a linearly

polarized, continuous light wave oscillating at a frequency, ω, that is given by the following equation:

$$\begin{aligned}\vec{E} &= E\hat{\epsilon}\cos(\hat{k}\cdot\hat{r}-\omega t) \\ &= \frac{E\hat{\epsilon}}{2}\left[e^{i(\hat{k}\cdot\hat{r}-\omega t)}+e^{-i(\hat{k}\cdot\hat{r}-\omega t)}\right]\end{aligned} \tag{A.3}$$

where E is the amplitude of the electric field, $\hat{\epsilon}$ is a unit vector specifying the direction of the electric field polarization, $\hat{k}$ is the propagation vector, and $\hat{r}$ is the position vector.

Eq. A.3 is further simplified by invoking the electric dipole approximation. This approximation assumes that the wavelength of the perturbing electromagnetic field is much larger than the dimensions of the molecules being excited. This means that the molecules being excited experience an electric field that is spatially constant, so that:

$$e^{\pm i(\hat{k}\cdot\hat{r})} = 1 \pm i\hat{k}\cdot\hat{r} \pm ... \approx 1 \tag{A.4}$$

$$\vec{E} \approx \frac{E\hat{\epsilon}}{2}\left[e^{-i\omega t}+e^{i\omega t}\right] \tag{A.5}$$

Given eq. A.5, $\hat{H}'(t)$ can be be written as the following:

$$\begin{aligned}\hat{H}'(t) &= -\frac{E_I}{2}\hat{\epsilon}_I\cdot\hat{\mu}e^{-i\omega_I t} - \frac{E_S}{2}\hat{\epsilon}_S\cdot\hat{\mu}e^{i\omega_S t} \\ &= Ve^{-i\omega_I t} + V^{\dagger}e^{i\omega_S t}\end{aligned} \tag{A.6}$$

where the subscripts I and S denote the incident and scattered photons. The operator, $\hat{\mu}$, is the transition dipole moment that couples the molecule to the perturbing radiation field.

We assume that the molecule starts in an initial state $|i\rangle$ and, following the perturbation, ends in state $|f\rangle$. The probability amplitude associated with this transition is:

$$\begin{aligned}\langle f|\hat{U}_I(t,0)|i\rangle &= \langle f|\hat{U}_I^{(0)}(t,0)|i\rangle + \langle f|\hat{U}_I^{(1)}(t,0)|i\rangle + \langle f|\hat{U}_I^{(2)}(t,0)|i\rangle \\ &= \delta_{fi} + \left(\frac{1}{i\hbar}\right)\langle f|\int_0^t dt_1 \hat{H}_I(t_1)\,|i\rangle + \left(\frac{1}{i\hbar}\right)^2 \langle f|\int_0^t dt_1 \int_0^{t_1} dt_2 \hat{H}_I(t_1)\hat{H}_I(t_2)\,|i\rangle\end{aligned} \tag{A.7}$$

The Kronecker delta, δ_{fi}, indicates that the $\langle f|\hat{U}_I^{(0)}(t,0)|i\rangle$ term is non-zero only if no transition occurs, (*i.e.*, $i = f$). In contrast, the $\langle f|\hat{U}_I^{(1)}(t,0)|i\rangle$ term represents one-photon absorption and emission processes. Thus, only the $\langle f|\hat{U}_I^{(2)}(t,0)|i\rangle$ term of eq. A.7 contributes to the probability amplitude associated with two photon scattering processes such as Rayleigh and Raman:

$$\langle f|\hat{U}_I^{(2)}(t,0)|i\rangle = \left(\frac{1}{i\hbar}\right)^2 \langle f|\int_0^t dt_1 \int_0^{t_1} dt_2 \hat{H}_I(t_1)\hat{H}_I(t_2)\,|i\rangle \tag{A.8}$$

To evaulate eq. A.8, we use the following identity:

$$\sum_r |r\rangle\langle r| = 1 \tag{A.9}$$

Substituting eqs. A.6 and A.9 into A.8, gives the following equation:

$$\begin{aligned}\langle f|\hat{U}_I^{(2)}(t,0)|i\rangle &= \left(\frac{1}{i\hbar}\right)^2 \sum_r \int_0^t dt_1 \langle f|\,\hat{H}_I(t_1)\,|r\rangle \int_0^{t_1} dt_2 \langle r|\,\hat{H}_I(t_2)\,|i\rangle \\ &= \left(\frac{1}{i\hbar}\right)^2 \sum_r \int_0^t dt_1 \langle f|\,e^{i\hat{H}_0 t_1/\hbar}[Ve^{-i\omega_I t_1} + V^\dagger e^{i\omega_S t_1}]e^{-i\hat{H}_0 t_1/\hbar}\,|r\rangle \\ &\times \int_0^{t_1} dt_2 \langle r|\,e^{i\hat{H}_0 t_2/\hbar}[Ve^{-i\omega_I t_2} + V^\dagger e^{i\omega_S t_2}]e^{-i\hat{H}_0 t_2/\hbar}\,|i\rangle \\ &= \left(\frac{1}{i\hbar}\right)^2 \sum_r \int_0^t dt_1 \Big[\langle f|V|r\rangle\, e^{i(\omega_{fr}-\omega_I)t_1} + \langle f|V^\dagger|r\rangle\, e^{i(\omega_{fr}+\omega_S)t_1}\Big] \\ &\times \int_0^{t_1} dt_2 \Big[\langle r|V|i\rangle\, e^{i(\omega_{ri}-\omega_I)t_2} + \langle r|V^\dagger|i\rangle\, e^{i(\omega_{ri}+\omega_S)t_2}\Big]\end{aligned} \tag{A.10}$$

where $\omega_{fr} = \omega_f - \omega_r$ and $\omega_{fr} = \omega_f - \omega_r$. Integrating eq. A.10 with respect to t_2 gives:

$$\left(\frac{1}{i\hbar}\right)^2 \sum_r \int_0^t dt_1 \Bigg[\frac{\langle f|V|r\rangle \langle r|V|i\rangle \, e^{i(\omega_{fi}-2\omega_I)t_1} - \langle f|V|r\rangle \langle r|V|i\rangle \, e^{i(\omega_{fr}-\omega_I)t_1}}{i(\omega_{ri}-\omega_I)} + \frac{\langle f|V|r\rangle \langle r|V^\dagger|i\rangle \, e^{i(\omega_{fi}-\omega_I+\omega_S)t_1} - \langle f|V|r\rangle \langle r|V^\dagger|i\rangle \, e^{i(\omega_{fr}-\omega_I)t_1}}{i(\omega_{ri}+\omega_S)} + \frac{\langle f|V^\dagger|r\rangle \langle r|V|i\rangle \, e^{i(\omega_{fi}-\omega_I+\omega_S)t_1} - \langle f|V^\dagger|r\rangle \langle r|V|i\rangle \, e^{i(\omega_{fr}+\omega_S)t_1}}{i(\omega_{ri}-\omega_I)} + \frac{\langle f|V^\dagger|r\rangle \langle r|V^\dagger|i\rangle \, e^{i(\omega_{fi}+2\omega_S)t_1} - \langle f|V^\dagger|r\rangle \langle r|V^\dagger|i\rangle \, e^{i(\omega_{fr}+\omega_S)t_1}}{i(\omega_{ri}+\omega_S)} \Bigg] \tag{A.11}$$

where we have utilized the fact that $\omega_{fi} = \omega_{fr} + \omega_{ri}$.

Expression A.11 is modulated in five frequencies: $\omega_{fi} - \omega_I + \omega_S$, $\omega_{fi} - 2\omega_I$, $\omega_{fi} + 2\omega_S$, $\omega_{fr} - \omega_I$, and $\omega_{fr} + \omega_S$. The terms containing $\omega_{fr} - \omega_I$ and $\omega_{fr} + \omega_S$ represent transitions that are non-resonant with the initial or final states. Integrating these terms with respect to t_1 averages to zero over short timescales. Thus, the contributions of these terms in A.11 can be neglected.

The term containing $\omega_{fi} - 2\omega_I$ corresponds to two-photon absorption (Figure 2.3d) since the frequency of the incident photon is half the frequency associated with the transition from the initial to the final state. Similarly, the term containing $\omega_{fi} + 2\omega_S$ corresponds to a two-photon emission (Figure 2.3e). The only relevant terms for Rayleigh and Raman scattering contain $\omega_{fi} - \omega_I + \omega_S$. Integrating the terms that contain $\omega_{fi} - \omega_I + \omega_S$ with respect to t_1 gives the following probability amplitude:

$$\begin{aligned}
\langle f|\hat{U}_I^{(2)}(t,0)|i\rangle &= \left(\frac{1}{i\hbar}\right)^2 \sum_r \int_0^t dt_1 \left[\frac{\langle f|V|r\rangle \langle r|V^\dagger|i\rangle \, e^{i(\omega_{fi}-\omega_I+\omega_S)t_1}}{i(\omega_{ri}+\omega_S)} + \frac{\langle f|V^\dagger|r\rangle \langle r|V|i\rangle \, e^{i(\omega_{fi}-\omega_I+\omega_S)t_1}}{i(\omega_{ri}-\omega_I)} \right] \\
&= \frac{1}{\hbar^2} \sum_r \left[\frac{\langle f|V^\dagger|r\rangle \langle r|V|i\rangle}{\omega_{ri}-\omega_I} + \frac{\langle f|V|r\rangle \langle r|V^\dagger|i\rangle}{\omega_{ri}+\omega_S} \right] \left[\frac{e^{i(\omega_{fi}-\omega_I+\omega_S)t}-1}{\omega_{fi}-\omega_I+\omega_S} \right] \\
&= \frac{E_L E_S}{4\hbar^2} \sum_r \left[\frac{\langle f|\hat{\epsilon}_S\cdot\hat{\mu}_\rho|r\rangle \langle r|\hat{\epsilon}_I\cdot\hat{\mu}_\sigma|i\rangle}{\omega_{ri}-\omega_I} + \frac{\langle f|\hat{\epsilon}_I\cdot\hat{\mu}_\sigma|r\rangle \langle r|\hat{\epsilon}_S\cdot\hat{\mu}_\rho|i\rangle}{\omega_{ri}+\omega_S} \right] \\
&\times \left[\frac{e^{i(\omega_{fi}-\omega_I+\omega_S)t}-1}{\omega_{fi}-\omega_I+\omega_S} \right]
\end{aligned} \tag{A.12}$$

where the subscripts ρ and σ represent the directions of the transition moments.

A.1.2 Fermi's Golden Rule and the Transition Scattering Rate

The transition probability, $P_{fi}(t)$, of finding the system in state $|f\rangle$ at time t is the square of eq. A.12:

$$\begin{aligned}
P_{i\to f}(t) &= |\,\langle f|U_I^{(2)}(t,0)|i\rangle\,|^2 \\
&= \frac{E_I^2 E_S^2}{16\hbar^4}\sum_r \left|\frac{\langle f|\hat{\epsilon}_S\cdot\hat{\mu}_\rho|r\rangle\,\langle r|\hat{\epsilon}_I\cdot\hat{\mu}_\sigma|i\rangle}{\omega_{ri}-\omega_I} + \frac{\langle f|\hat{\epsilon}_I\cdot\hat{\mu}_\sigma|r\rangle\,\langle r|\hat{\epsilon}_S\cdot\hat{\mu}_\rho|i\rangle}{\omega_{ri}+\omega_S}\right|^2 \\
&\times \left|\frac{e^{i(\omega_{fi}-\omega_I+\omega_S)t}-1}{\omega_{fi}-\omega_I+\omega_S}\right|^2 \\
&= \frac{E_I^2 E_S^2}{16\hbar^4}\sum_r \left|\frac{\langle f|\hat{\epsilon}_S\cdot\hat{\mu}_\rho|r\rangle\,\langle r|\hat{\epsilon}_I\cdot\hat{\mu}_\sigma|i\rangle}{\omega_{ri}-\omega_I} + \frac{\langle f|\hat{\epsilon}_I\cdot\hat{\mu}_\sigma|r\rangle\,\langle r|\hat{\epsilon}_S\cdot\hat{\mu}_\rho|i\rangle}{\omega_{ri}+\omega_S}\right|^2 \\
&\times \left[\frac{4\sin^2[(\omega_{fi}-\omega_I+\omega_S)t/2]}{(\omega_{fi}-\omega_I+\omega_S)^2}\right]
\end{aligned} \tag{A.13}$$

Eq. A.13 is the probability of a transition to occur between two discrete states, $|i\rangle$ and $|f\rangle$. However, we assume a central tenant of Fermi's Golden Rule, *viz.* that the molecule is being excited into a continuum of $|f\rangle$ states that span a range of frequencies. If we assume that, in this continuum, each $|f\rangle$ state is independent of the others, then the total transition probability is simply the sum of the individual transition probabilities:

$$\begin{aligned}
P(t) &= \sum_f P_{i\to f}(t) \\
&= \sum_f \frac{E_I^2 E_S^2}{16\hbar^4}\sum_r \left|\frac{\langle f|\hat{\epsilon}_S\cdot\hat{\mu}_\rho|r\rangle\,\langle r|\hat{\epsilon}_I\cdot\hat{\mu}_\sigma|i\rangle}{\omega_{ri}-\omega_I} + \frac{\langle f|\hat{\epsilon}_I\cdot\hat{\mu}_\sigma|r\rangle\,\langle r|\hat{\epsilon}_S\cdot\hat{\mu}_\rho|i\rangle}{\omega_{ri}+\omega_S}\right|^2 \\
&\times \left[\frac{4\sin^2[(\omega_{fi}-\omega_I+\omega_S)t/2]}{(\omega_{fi}-\omega_I+\omega_S)^2}\right]
\end{aligned} \tag{A.14}$$

The index, f, is continuous. To convert the summation to an integral, we introduce ρ_f, the density of states, which is defined as the number of levels per unit energy:

$$\rho_f = \frac{dn_f}{dE_f} = \frac{1}{\hbar}\frac{dn_f}{d\omega_f} \tag{A.15}$$

where ρ_f is assumed to be a continuous function over $\hbar\omega_f$. Substituting eq. A.15 into A.14 gives the following:

$$P(t) = \frac{E_I^2 E_S^2}{16\hbar^3} \int_{\omega_i-\Delta\omega/2}^{\omega_i+\Delta\omega/2} d\omega_f \rho_f \sum_r \left| \frac{\langle f|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|r\rangle \langle r|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|r\rangle \langle r|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right|^2 \times \left[\frac{4\sin^2[(\omega_{fi} - \omega_I + \omega_S)t/2]}{(\omega_{fi} - \omega_I + \omega_S)^2} \right] \tag{A.16}$$

Further simplifying assumptions are necessary in order to integrate eq. A.16. Although ρ_f and the transition moment integrals are functions of ω_f, we assume that they vary slowly, and thus are approximately constant over the range of frequencies being considered. This means that eq. A.16 can be written as:

$$P(t) = \frac{E_I^2 E_S^2}{16\hbar^3} \rho_f \sum_r \left| \frac{\langle f|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|r\rangle \langle r|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|r\rangle \langle r|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right|^2 \times \int_{\omega_i-\Delta\omega/2}^{\omega_i+\Delta\omega/2} d\omega_f \left[\frac{4\sin^2[(\omega_{fi} - \omega_I + \omega_S)t/2]}{(\omega_{fi} - \omega_I + \omega_S)^2} \right] \tag{A.17}$$

To integrate eq. A.17, the following substitution is made:

$$u = (\omega_{fi} - \omega_I + \omega_S)t/2 \tag{A.18}$$

$$P(t) = \frac{E_I^2 E_S^2}{8\hbar^3} \rho_f t \sum_r \left| \frac{\langle f|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|r\rangle \langle r|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|r\rangle \langle r|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right|^2 \times \int_{-\Delta\omega t/4}^{\Delta\omega t/4} du \left[\frac{\sin^2(u)}{u^2} \right] \tag{A.19}$$

The limits of integration in eq. A.19 can be extended to $\pm\infty$ without serious error so that:

$$\begin{aligned} P(t) &= \frac{E_I^2 E_S^2}{8\hbar^3} \rho_f t \sum_r \left| \frac{\langle f|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|r\rangle \langle r|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|r\rangle \langle r|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right|^2 \\ &\times \int_{-\infty}^{+\infty} du \left[\frac{\sin^2(u)}{u^2} \right] \\ &= \frac{\pi E_I^2 E_S^2}{8\hbar^3} \sum_r \left| \frac{\langle f|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|r\rangle \langle r|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|r\rangle \langle r|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right|^2 \rho_f t \end{aligned} \tag{A.20}$$

Eq. A.20 indicates that the total transition probability is linearly proportional to time, so that the scattering rate, W_{fi}, can be defined as follows:

$$\begin{aligned} W_{fi} &= \frac{dP(t)}{dt} \\ &= \frac{\pi E_I^2 E_S^2}{8\hbar^3} \sum_r \left| \frac{\langle f|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|r\rangle \langle r|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{\epsilon}_I \cdot \hat{\mu}_\sigma|r\rangle \langle r|\hat{\epsilon}_S \cdot \hat{\mu}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right|^2 \rho_f \end{aligned} \tag{A.21}$$

A.2 DERIVATION OF THE ALBRECHT *A*, *B*, AND *C* TERMS

In this section, the Albrecht A, B, and C terms (eqs. 2.16–2.18) are derived. We begin by using eqs. 2.11 and 2.12 to write the relevant electronic transition moment integrals of eq. 2.10 as:

$$\begin{aligned} \langle \phi_g|\hat{r}_\rho|\phi_e\rangle \approx \langle \phi_g^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle &+ \frac{1}{\hbar} \sum_{a,s\neq e} \frac{(h_a^{(0)})_{es}}{\omega_{es}^0} Q_a \langle \phi_g^{(0)}|\hat{r}_\rho|\phi_s^{(0)}\rangle \\ &+ \frac{1}{\hbar} \sum_{a,t\neq g} \frac{(h_a^{(0)})_{gt}}{\omega_{gt}^0} Q_a \langle \phi_t^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle \end{aligned} \tag{A.22}$$

$$\begin{aligned} \langle \phi_e|\hat{r}_\sigma|\phi_g\rangle \approx \langle \phi_e^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle &+ \frac{1}{\hbar} \sum_{a,t\neq g} \frac{(h_a^{(0)})_{gt}}{\omega_{gt}^0} Q_a \langle \phi_e^{(0)}|\hat{r}_\sigma|\phi_t^{(0)}\rangle \\ &+ \frac{1}{\hbar} \sum_{a,s\neq e} \frac{(h_a^{(0)})_{es}}{\omega_{es}^0} Q_a \langle \phi_s^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle \end{aligned} \tag{A.23}$$

$$\begin{aligned} \langle \phi_g|\hat{r}_\sigma|\phi_e\rangle \approx \langle \phi_g^{(0)}|\hat{r}_\sigma|\phi_e^{(0)}\rangle &+ \frac{1}{\hbar} \sum_{a,s\neq e} \frac{(h_a^{(0)})_{es}}{\omega_{es}^0} Q_a \langle \phi_g^{(0)}|\hat{r}_\sigma|\phi_s^{(0)}\rangle \\ &+ \frac{1}{\hbar} \sum_{a,t\neq g} \frac{(h_a^{(0)})_{gt}}{\omega_{gt}^0} Q_a \langle \phi_t^{(0)}|\hat{r}_\sigma|\phi_e^{(0)}\rangle \end{aligned} \tag{A.24}$$

$$\langle\phi_e|\hat{r}_\rho|\phi_g\rangle \approx \langle\phi_e^{(0)}|\hat{r}_\rho|\phi_g^{(0)}\rangle + \frac{1}{\hbar}\sum_{a,t\neq g}\frac{(h_a^{(0)})_{gt}}{\omega_{gt}^0}Q_a\,\langle\phi_e^{(0)}|\hat{r}_\rho|\phi_t^{(0)}\rangle$$
$$+\frac{1}{\hbar}\sum_{a,s\neq e}\frac{(h_a^{(0)})_{es}}{\omega_{es}^0}Q_a\,\langle\phi_s^{(0)}|\hat{r}_\rho|\phi_g^{(0)}\rangle \tag{A.25}$$

These expressions can be utilized to write the products of the vibronic transition moment integrals in eq. 2.10 as:

$$\langle\chi_{gf'}|\,\langle\phi_g|\hat{r}_\rho|\phi_e\rangle\,|\chi_{er'}\rangle\,\langle\chi_{er'}|\,\langle\phi_e|\hat{r}_\sigma|\phi_g\rangle\,|\chi_{gi'}\rangle =$$
$$\Bigg[\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\chi_{gf'}|\chi_{er'}\rangle + \frac{1}{\hbar}\sum_{a,s\neq e}\frac{(h_a^{(0)})_{es}}{\omega_{es}^0}\,\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_s^{(0)}\rangle\,\langle\chi_{gf'}|Q_a|\chi_{er'}\rangle$$
$$\frac{1}{\hbar}\sum_{a,t\neq g}\frac{(h_a^{(0)})_{gt}}{\omega_{gt}^0}\,\langle\phi_t^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\chi_{gf'}|Q_a|\chi_{er'}\rangle\Bigg]\Bigg[\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle\,\langle\chi_{er'}|\chi_{gi'}\rangle$$
$$+\frac{1}{\hbar}\sum_{a,t\neq g}\frac{(h_a^{(0)})_{gt}}{\omega_{gt}^0}\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_t^{(0)}\rangle\,\langle\chi_{er'}|Q_a|\chi_{gi'}\rangle + \frac{1}{\hbar}\sum_{a,s\neq e}\frac{(h_a^{(0)})_{es}}{\omega_{es}^0}\,\langle\phi_s^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle\,\langle\chi_{er'}|Q_a|\chi_{gi'}\rangle\Bigg]$$
$$\approx \langle\phi_g|\hat{r}_\rho|\phi_e\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle\,\langle\chi_{gf'}|\chi_{er'}\rangle\,\langle\chi_{er'}|\chi_{gi'}\rangle +$$
$$+\frac{1}{\hbar}\sum_{a,t\neq g}\frac{\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_t^{(0)}\rangle\,\langle\phi_t^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_g^{(0)}\rangle}{\omega_g^0-\omega_t^0}\,\langle\chi_{gf'}|\chi_{er'}\rangle\,\langle\chi_{er'}|Q_a|\chi_{gi'}\rangle$$
$$+\frac{1}{\hbar}\sum_{a,s\neq e}\frac{\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_s^{(0)}\rangle\,\langle\phi_s^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle}{\omega_e^0-\omega_s^0}\,\langle\chi_{gf'}|\chi_{er'}\rangle\,\langle\chi_{er'}|Q_a|\chi_{gi'}\rangle$$
$$+\frac{1}{\hbar}\sum_{a,s\neq e}\frac{\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_s^{(0)}\rangle\,\langle\phi_s^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle}{\omega_e^0-\omega_s^0}\,\langle\chi_{gf'}|Q_a|\chi_{er'}\rangle\,\langle\chi_{er'}|\chi_{gi'}\rangle$$
$$+\frac{1}{\hbar}\sum_{a,t\neq g}\frac{\langle\phi_g^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_t^{(0)}\rangle\,\langle\phi_t^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle}{\omega_e^0-\omega_s^0}\,\langle\chi_{gf'}|Q_a|\chi_{er'}\rangle\,\langle\chi_{er'}|\chi_{gi'}\rangle \tag{A.26}$$

$$
\begin{aligned}
&\langle\chi_{gf'}|\,\langle\phi_g|\hat{r}_\sigma|\phi_e\rangle\,|\chi_{er'}\rangle\,\langle\chi_{er'}|\,\langle\phi_e|\hat{r}_\sigma|\phi_g\rangle\,|\chi_{gi'}\rangle = \cdots \approx \\
&\langle\phi_g|\hat{r}_\rho|\phi_e\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle\,\langle\chi_{gf'}|\chi_{er'}\rangle\,\langle\chi_{er'}|\chi_{gi'}\rangle + \\
&+\frac{1}{\hbar}\sum_{a,t\neq g}\frac{\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_t^{(0)}\rangle\,\langle\phi_t^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_g^{(0)}\rangle}{\omega_g^0-\omega_t^0}\,\langle\chi_{gf'}|\chi_{er'}\rangle\,\langle\chi_{er'}|Q_a|\chi_{gi'}\rangle \\
&+\frac{1}{\hbar}\sum_{a,s\neq e}\frac{\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_s^{(0)}\rangle\,\langle\phi_s^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle}{\omega_e^0-\omega_s^0}\,\langle\chi_{gf'}|\chi_{er'}\rangle\,\langle\chi_{er'}|Q_a|\chi_{gi'}\rangle \\
&+\frac{1}{\hbar}\sum_{a,s\neq e}\frac{\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_s^{(0)}\rangle\,\langle\phi_s^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle}{\omega_e^0-\omega_s^0}\,\langle\chi_{gf'}|Q_a|\chi_{er'}\rangle\,\langle\chi_{er'}|\chi_{gi'}\rangle \\
&+\frac{1}{\hbar}\sum_{a,t\neq g}\frac{\langle\phi_g^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_t^{(0)}\rangle\,\langle\phi_t^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle}{\omega_e^0-\omega_s^0}\,\langle\chi_{gf'}|Q_a|\chi_{er'}\rangle\,\langle\chi_{er'}|\chi_{gi'}\rangle
\end{aligned}
\tag{A.27}
$$

Eqs. A.26 and A.27 can be substituted into eq. 2.10 to derive eq. 2.15:

$$
\begin{aligned}
(\alpha_{\rho\sigma})_{gf',gi'} &= \frac{1}{\hbar}\sum_{e\neq g}\sum_{r'}\left[\frac{\langle\phi_g|\hat{r}_\rho|\phi_e\rangle\,\langle\phi_e|\hat{r}_\sigma|\phi_g\rangle}{\omega_{er'}-\omega_{gi'}-\omega_I}+\frac{\langle\phi_g|\hat{r}_\sigma|\phi_e\rangle\,\langle\phi_e|\hat{r}_\rho|\phi_g\rangle}{\omega_{er'}-\omega_{gi'}+\omega_S}\right]\langle\chi_{gf'}|\chi_{er'}\rangle\,\langle\chi_{er'}|\chi_{gi'}\rangle \\
&+\frac{1}{\hbar^2}\sum_{e\neq g}\sum_{r'}\sum_{a}\sum_{s\neq e}\left\{\left[\frac{\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_s^{(0)}\rangle\,\langle\phi_s^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle}{\omega_{er'}-\omega_{gi'}-\omega_I}\right.\right. \\
&\left.+\frac{\langle\phi_g^{(0)}|\hat{r}_\sigma|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_s^{(0)}\rangle\,\langle\phi_s^{(0)}|\hat{r}_\rho|\phi_g^{(0)}\rangle}{\omega_{er'}-\omega_{gi'}+\omega_S}\right]\left[\frac{\langle\chi_{gf'}|\chi_{er'}\rangle\,\langle\chi_{er'}|\hat{Q}_a|\chi_{gi'}\rangle}{\omega_e^{(0)}-\omega_s^{(0)}}\right] \\
&+\left[\frac{\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_s^{(0)}\rangle\,\langle\phi_s^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle}{\omega_{er'}-\omega_{gi'}-\omega_I}\right. \\
&\left.\left.+\frac{\langle\phi_g^{(0)}|\hat{r}_\sigma|\phi_s^{(0)}\rangle\,\langle\phi_s^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\rho|\phi_g^{(0)}\rangle}{\omega_{er'}-\omega_{gi'}+\omega_S}\right]\left[\frac{\langle\chi_{gf'}|\hat{Q}_a|\chi_{er'}\rangle\,\langle\chi_{er'}|\chi_{gi'}\rangle}{\omega_e^{(0)}-\omega_s^{(0)}}\right]\right\} \\
&+\frac{1}{\hbar^2}\sum_{e\neq g}\sum_{r'}\sum_{a}\sum_{t\neq e}\left\{\left[\frac{\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_t^{(0)}\rangle\,\langle\phi_t^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_g^{(0)}\rangle}{\omega_{er'}-\omega_{gi'}-\omega_I}\right.\right. \\
&\left.+\frac{\langle\phi_g^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_t^{(0)}\rangle\,\langle\phi_t^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_g^{(0)}\rangle}{\omega_{er'}-\omega_{gi'}+\omega_S}\right]\left[\frac{\langle\chi_{gf'}|\chi_{er'}\rangle\,\langle\chi_{er'}|\hat{Q}_a|\chi_{gi'}\rangle}{\omega_e^{(0)}-\omega_s^{(0)}}\right] \\
&+\left[\frac{\langle\phi_g^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_t^{(0)}\rangle\,\langle\phi_t^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle}{\omega_{er'}-\omega_{gi'}-\omega_I}\right. \\
&\left.\left.+\left[\frac{\langle\phi_g^{(0)}|(\partial\hat{H}/\partial Q_a)_0|\phi_t^{(0)}\rangle\,\langle\phi_t^{(0)}|\hat{r}_\rho|\phi_e^{(0)}\rangle\,\langle\phi_e^{(0)}|\hat{r}_\sigma|\phi_g^{(0)}\rangle}{\omega_{er'}-\omega_{gi'}+\omega_S}\right]\left[\frac{\langle\chi_{gf'}|\hat{Q}_a|\chi_{er'}\rangle\,\langle\chi_{er'}|\chi_{gi'}\rangle}{\omega_g^{(0)}-\omega_t^{(0)}}\right]\right\} \\
&= A+B+C
\end{aligned}
\tag{A.28}
$$

where the A, B, and C terms are defined by eqs. 2.16–2.18.

APPENDIX B

SUPPORTING INFORMATION FOR CHAPTER 3.0

B.1 X-RAY CRYSTAL STRUCTURE OF PROPANAMIDE

X-ray crystal diffraction was performed using a Bruker X8 Prospector Ultra equipped with a copper micro-focus tube (λ = 1.541 78 Å). A propanamide crystal specimen with approximate dimensions of 0.020 mm × 0.090 mm × 0.110 mm was used for structure determination. The crystal was mounted and placed in a cold N_2 stream (∼230 K) for data collection. Crystals were not well-formed, showing diffuse diffraction spots; however, we were able to determine a unit cell and solve a crystal structure.

The frames were integrated with the Bruker SAINT software package. The integration of the data using a monoclinic unit cell yielded a total of 2212 reflections to a maximum θ angle of 68.09° (0.83 Å resolution), of which 689 were independent (average redundancy 3.210, completeness = 83.5%, R_{int} = 8.52%, R_{sig} = 7.03%) and 364 (52.83%) were greater than $2\sigma(F^2)$.

The final cell constants (Table B1) of a = 8.851(4) Å, b = 5.750(2) Å, c = 9.766(3) Å, β = 114.780(15)°, volume = 451.3(3) $Å^3$, are based upon the refinement of the XYZ-centroids of reflections above 20 σ(I). Unit-cell parameters and analysis of systematic absences indicated propanamide crystallized in a $P2_1/c$ space group.

The structure (Figure B1) was solved via direct methods, which located all of the non-hydrogen atoms. Idealized atom positions were calculated for all hydrogen atoms, except for NH_2 hydrogen atoms (see Table B2–Table B6). The final anisotropic least-squares refinement

on F^2 converged at R1 = 9.53%, for the observed data and wR2 = 24.72% for all data, and the goodness-of-fit was 1.304. The final Fourier map contained no significant residual electron density.

B.2 DECONVOLUTION OF UVRR SPECTRA

UVRR spectra of propanamide in different acetonitrile and water mixtures were deconvoluted into a sum of a minumum number of Gaussian or Lorentzian bands on a linear background using the GRAMS/AI 8.0 software suite (Thermo Fisher). Figure B2 and Figure B3 show the deconvolution of the 204 nm excitation UVRR spectra of propanamide in pure acetonitrile and pure water.

Table B1: Summary of Crystallographic Data for $CH_3CH_2CONH_2$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

Molecular formula	C_3H_7NO
Temperature (K)	230(2)
Wavelength (Å)	1.54178
Theta range (°)	9.18-68.09
Cell setting	monoclinic
Space group	$P2_1/c$
a (Å)	8.851(4)
b (Å)	5.750(2)
c (Å)	9.766(3)
α (°)	90
β (°)	114.780(15)
γ (°)	90
V (Å^3)	451.259
Z	4
Calc. density (g·cm^{-1})	1.076
R1	0.0953
wR2	0.2472

Table B2: Atomic Coordinates and Equivalent Isotropic Displacement Parameters (Å). Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

	x/a	y/b	z/c	U(eq)
O9	0.6355(5)	0.7687(5)	0.6186(3)	0.1029(15)
N10	0.6057(7)	0.8705(7)	0.3890(4)	0.0958(16)
C1	0.8330(8)	0.3790(10)	0.6341(7)	0.120(2)
C2	0.7834(9)	0.5471(11)	0.5093(6)	0.121(2)
C6	0.6678(7)	0.7373(7)	0.5094(5)	0.0867(16)

Table B3: Anisotropic Atomic Displacement Parameters ($Å^2$). Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
O9	0.166(3)	0.095(2)	0.0723(19)	0.0141(14)	0.075(2)	0.0239(18)
N10	0.159(4)	0.084(3)	0.067(2)	0.004(2)	0.069(3)	0.014(3)
C1	0.152(6)	0.092(3)	0.116(4)	0.007(3)	0.055(4)	0.018(3)
C2	0.169(6)	0.121(4)	0.098(4)	0.022(3)	0.080(4)	0.047(4)
C6	0.135(4)	0.072(3)	0.070(2)	0.002(2)	0.060(3)	0.004(2)

Table B4: Comparison of Bond Lengths (Å) Between the DFT-optimized and X-ray Crystal Structure. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

	Crys. Struc.	DFT Calc.
r(C6=O9)	1.228(6)	1.223
r(C6-N10)	1.315(5)	1.353
r(C6-C2)	1.498(8)	1.519
r(C2-C1)	1.470(8)	1.521
r(C1-H3)	0.97	1.090
r(C1-H4)	0.97	1.091
r(C1-H5)	0.97	1.091
r(C2-H7)	0.98	1.095
r(C2-H8)	0.98	1.093
r(N10-H11)	0.83(5)	1.007
r(N10-H12)	0.97(6)	1.008

Table B5: Comparison of Torsion Angles (°) Between the DFT-optimized and X-ray Crystal Structure. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

	Crys. Struc.	DFT Calc.
τ(N10C6C2C1)	-171.5(5)	169.60
τ(O9C6C2C1)	9.8(7)	-10.89
τ(H12N10C6O9)	-16(4)	-2.69
τ(H12N10C6C2)	166(4)	176.83
τ(H11N10C6O9)	-176(3)	-178.79
τ(H11N10C6C2)	5(3)	0.72
τ(H5C1C2H8)	-56.4	-58.32
τ(H5C1C2H7)	59.5	59.54
τ(H5C1C2C6)	-178.4	179.58
τ(H3C1C2H8)	-176.4	-178.37
τ(H3C1C2H7)	-60.4	-60.51
τ(H3C1C2C6)	61.6	59.54
τ(H4C1C2H8)	63.6	61.64
τ(H4C1C2H7)	179.6	179.50
τ(H4C1C2C6)	-58.4	-60.46
τ(H8C2C6O9)	-112.2	-134.50
τ(H8C2C6N10)	66.5	-45.99

Table B6: Comparison of Bond Angles (°) Between the DFT-optimized and X-ray Crystal Structure. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

	Crys. Struc.	DFT Calc.
θ(H12N10H11)	121(4)	118.56
θ(H12N10C6)	118(3)	119.60
θ(H11N10C6)	118(3)	121.73
θ(H5C1H3)	109.4	108.43
θ(H5C1H4)	109.5	108.32
θ(H5C1C2)	109.5	110.17
θ(H3C1H4)	109.5	107.97
θ(H3C1C2)	109.5	110.90
θ(H4C1C2)	109.5	110.17
θ(C1C2H8)	108.2	111.15
θ(C1C2H7)	108.2	110.72
θ(C1C2C6)	116.5(5)	112.88
θ(H8C2H7)	107.3	106.26
θ(H8C2C6)	108.2	108.44
θ(H7C2C6)	108.2	107.09
θ(O9C6N10)	121.9(4)	122.07
θ(O9C6C2)	121.1(4)	122.91
θ(N10C6C2)	117.0(4)	115.02

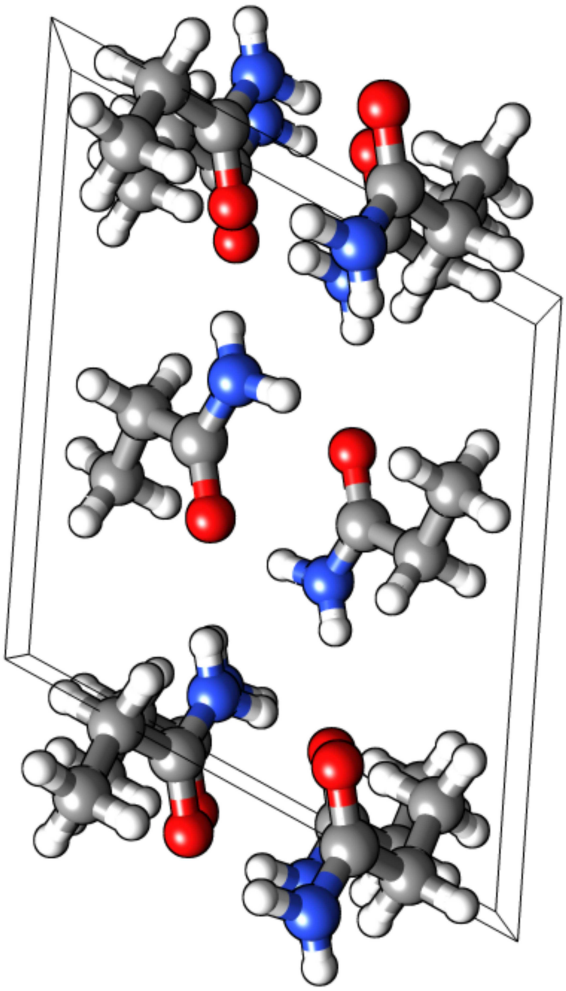


Figure B1: Crystal structure of propanamide. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

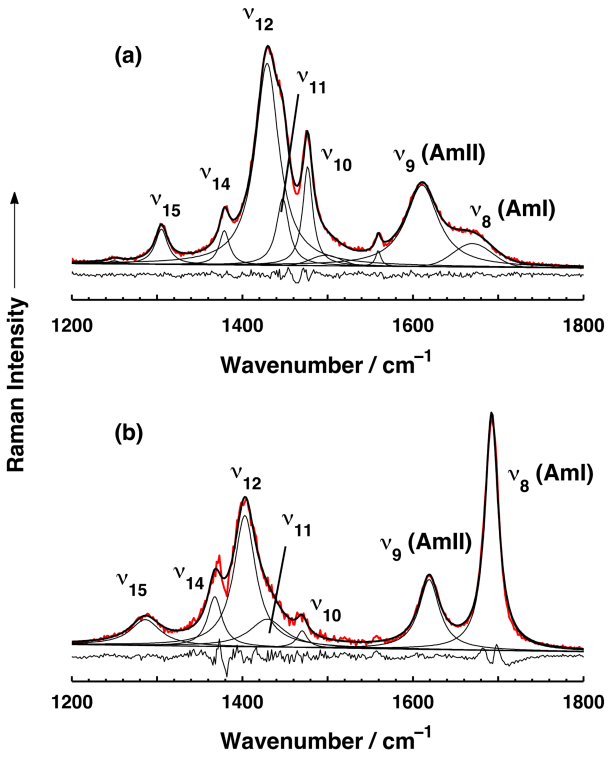


Figure B2: Spectral deconvolution of the 204 nm UVRR spectrum of propanamide in (a) water and (b) acetonitrile in the region from 1200–1800 cm^{-1}. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

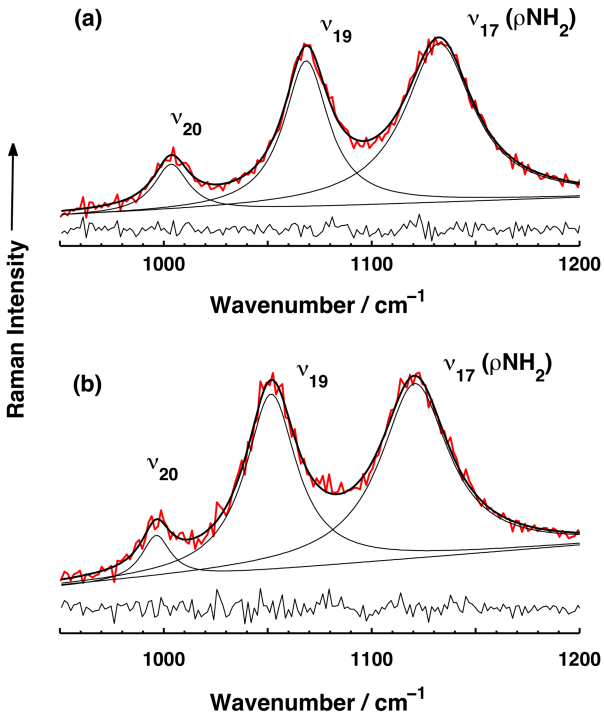


Figure B3: Spectral deconvolution of the 204 nm UVRR spectrum of propanamide in (a) water and (b) acetonitrile in the region from 800–1200 cm^{-1}. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

APPENDIX C

SUPPORTING INFORMATION FOR CHAPTER 4.0

C.1 DEPENDENCE OF THE AMIII^{P} BAND FREQUENCY ON THE OCCC DIHEDRAL ANGLE OF BUTYRAMIDE

In order to understand the asymmetry of the AmIII^{P} frequency dependence on the χ_3 dihedral angle of L-Gln, we investigated the achiral molecule butyramide, which is a model of the L-Gln side chain. DFT calculations were performed on butyramide using the same methods as described in the Computational Details section for L-Gln. Figure C1 shows the cosinuisoidal dependence of the AmIII^{P} band frequency on the OCCC dihedral angle of butyramide. Unlike L-Gln, there is no asymmetry in the points about 0°, and the data can be satisfactorily fit to the following equation:

$$\nu(\chi_3) = 1081 \text{ cm}^{-1} + (16 \text{ cm}^{-1}) \cos(2\chi_3) \tag{C.1}$$

C.2 X-RAY DIFFRACTION OF GLUTAMINE AND DERIVATIVES

We determined the unit cells and lattice constants of the following molecules: L-Gln, D-Gln, NAcGln, Gly-Gln, and Ser-Asn. We solved for the unit cells of each specimen by refining the XYZ-centroids of the reflections above 20σ(I). Information on the unit cells of each molecule are found in Table C1. All of these compounds show crystal lattice constants

that were essentially the same as those found in the Cambridge Crystallographic database [188, 217–219].

Also, we determined the crystal structures of all specimens examined, except for NAcGln. The crystal structures and atomic labeling schemes used in Tables S8-S13 are shown in Figure C2. All structures were solved and refined using the Bruker SHELXTL software package [220]. With the exception of GlnTBE, all the crystal structures that we determined match the Cambridge Crystallographic database entries.

The crystal structure of GlnTBE does not exist in the Cambridge Crystallographic database. The details regarding the data collection and structure refinement for GlnTBE are listed in Table C2 and Table C3. The integration of the data using a monoclinic unit cell yielded a total of 12444 reflections, of which 2246 were independent (average redundancy 5.541, completeness = 99.8%, R_{int} = 5.32%, R_{sig} = 5.00%), to a maximum θ angle of 68.25° (0.83 Å). There were 2181 independent reflections (∼97%) that were greater than $2\sigma(F^2)$. The intensities were corrected for absorption effects using the multi-scan method, SADABS. The calculated minimum and maximum transmission coefficients (based on the crystal size) for GlnTBE were 0.7800 and 0.8800, respectively. The structure of GlnTBE was solved using the P 1 2_1 1 space group with Z = 2 for the formula unit $C_9H_{18}N_2O_3$ HCl. The final anisotropic full-matrix least-squares refinement on F^2 with 161 variables converged at R1 = 6.09% for the observed data and wR2 = 15.75% for all data. The goodness-of-fit was 1.812. The largest peak in the final difference electron density synthesis was 1.207 e·Å^{-3} and the largest hole was -0.585 e·Å^{-3}. The calculated density, on the basis of the final model, was 1.239 g·cm^{-3} and F(000) was 256 e$^-$. Table C2 to Table C7 list the atomic coordinates, equivalent isotropic displacement parameters, anisotropic displacement parameters, bond lengths, and bond angles for the atoms in GlnTBE. The atomic labeling scheme of GlnTBE used for Table C2 to Table C7 is shown in Figure C3.

C.3 RAMAN BAND ASSIGNMENTS OF CRYSTALLINE GLUTAMINE AND DERIVATIVES

C.3.1 Spectral Deconvolution

Figure C4 and Figure C5 show the 633 nm excitation Raman and the 229 nm UVRR spectra, respectively, for each of the crystals examined. We used the GRAMS software suite (Version 8.0, Thermo Fisher Scientific, Inc.) to peak fit the spectra in order to locate bands in the region from ∼950 cm^{-1} to 1200 cm^{-1}. We modeled the visible Raman spectra shown in Figure C4 as a minimum sum of Lorentzian bands.

We modeled the UVRR spectra, shown in Figure C5, as a minimum sum of Voigt bands in order to account for the spectrometer transfer function. We determined the transfer function by illuminating the CCD camera with the 229 nm line from our frequency doubled Ar^+ laser. For the slit widths used in our measurements (∼50 µm), the spectrometer transfer function was well modeled by a Gaussian with a full-width-half-height (FWHH) of ∼2.3 cm^{-1}. We fixed the FWHH of the Gaussian component of the Voigt function to ∼2.3 cm^{-1} and allowed the Lorenztian width to vary. We found the average half-width-half height of the $AmIII^P$ vibration to be ∼6.6±2.4 cm^{-1}. Thus, we assume that this value is the homogeneous linewidth, Γ, of the $AmIII^P$ vibration.

C.3.2 Band Assignments

Table C8–Table C13 show our band assignments of the crystals examined. We used DFT calculations to aid in our band assignments. We also measured the spectra of N-deuterated crystals (data not shown) to verify the band assignments of any vibrations that contain NH_3 or NH_2 rocking components. The $AmIII^P$ band shifts upon N-deuteration since the vibration contains significant NH_2 rocking.

C.3.2.1 L-glutamine and D-glutamine L-Gln and D-Gln have essentially identical Raman spectra. The experimentally observed and calculated Raman frequencies are shown in Table C8 and Table C9 for L-Gln and D-Gln, respectively. We assign the ∼1205 cm^{-1} band to

the $C_\beta H_2$ twisting/C_α–C_β stretching vibration, and the ∼1166 cm^{-1} band is assigned to a C_α–H rocking/$C_\gamma H_2$ twisting vibration. The ∼1135 cm^{-1} and ∼1105 cm^{-1} bands exhibit large shifts in their N-deuterated spectra and are therefore assigned to C_β–C_γ stretching/$N_\epsilon H_2$ rocking and NH_3 rocking vibrations, respectively. The $AmIII^P$ vibration is assigned to the ∼1097 cm^{-1} band since this band shows a large shift upon N-deuteration. The remaining peaks are assigned to a C_α–N stretching mode (∼1086 cm^{-1}), NH_3 rocking/C_α–N stretching mode (∼1052 cm^{-1}), and a NH_3 rocking/C_α–C_β stretching mode (∼1000 cm^{-1}).

C.3.2.2 N-Acetyl-L-glutamine The band frequencies and assignments for NAcGln are found in Table C10. The ∼1180 cm^{-1} Raman band is assigned to the C_α–N stretching mode, which agrees with that of Kausar *et al.*'s assignment for a similar band observed in N-Acetyl-L-glutamic acid [221]. The ∼1138 cm^{-1} band is assigned to C_α–C_β stretching/C_α–N stretching mode, while the ∼1111 cm^{-1} band is assigned to a $N_\epsilon H_2$ rocking vibration. The assignment of the 1111 cm^{-1} band to a primary amide $N_\epsilon H_2$ rocking mode is consistent with the absence of a peak in this region in N-Acetyl-L-glutamic acid [221]. The ∼1071 cm^{-1} band, which appears as a shoulder feature, is assigned to the $AmIII^P$ vibration. The bands at ∼1061 cm^{-1} and ∼1022 cm^{-1} are assigned to CH_3 and $C_\beta H_2$ rocking modes, respectively. The ∼997 cm^{-1} band is assigned to a C–C stretching vibration of the acetyl group.

C.3.2.3 L-glutamine t-butyl ester HCl Table C11 displays the frequencies and assignments of crystalline GlnTBE. The ∼1195 cm^{-1} band is assigned to a C–O stretching/CH_3 rocking mode of the butyl ester group. Most of the remaining bands in the spectra contain significant $N_\epsilon H_2$ or NH_3 character since they shift upon N-deuteration. However, the ∼1043 cm^{-1} and ∼1030 cm^{-1} bands do not shift appreciably upon N-deuteration and are therefore assigned to C_α–N stretching, and CH_3 rocking vibrations, respectively. The bands located at ∼1117 cm^{-1} and ∼1105 cm^{-1} are assigned to the NH_3 rocking/C_β–C_γ stretching and NH_3 rocking/$N_\epsilon H_2$ rocking vibrations, respectively. We assign the ∼1151 cm^{-1} band to a C_α–H rocking/NH_3 rocking mode. The band located at ∼1082 cm^{-1} is assigned to the $AmIII^P$ vibration due to its sensitivity to N-deuteration. The ∼998 cm^{-1} band is assigned to a NH_3 rocking vibration.

C.3.2.4 Glycl-L-glutamine Band assignments of crystalline Gly-Gln are shown in Table C12. The band observed at ∼1124 cm^{-1} is assigned to NH_3 rocking mode, while the ∼1093 cm^{-1} band is assigned to the $AmIII^P$ vibration. We do not see the NH_3 rocking mode calculated to be at ∼1105 cm^{-1}. The remaining band assignments are shown in Table C12.

C.3.2.5 L-seryl-L-asparagine Table C13 shows the band assignments of crystalline Ser-Asn. The ∼1188 cm^{-1} band is assigned to a $C_{\beta'}H_2$ rocking/NH_3 rocking vibration due to its shift upon N-deuteration. The ∼1159 cm^{-1} band is assigned to a C_α–N stretching vibration, while the bands at ∼1121 cm^{-1} and ∼1108 cm^{-1} are both assigned to $C_{\beta'}$–$O_{\gamma'}H$ stretching modes of the serine side chain. We assign the $AmIII^P$ vibration to the ∼1051 cm^{-1} band. The remaining bands located at ∼1083 cm^{-1}, ∼1065 cm^{-1}, ∼1019 cm^{-1}, and ∼1004 cm^{-1} are assigned to $N_\delta H_2$ rocking, NH_3 rocking, $C_\beta H_2$ rocking, and $C_{\beta'}H_2$ rocking/NH_3 rocking modes, respectively.

C.4 DEPENDENCE OF THE $AMIII^P$ FREQUENCY ON HYDROGEN BONDING AND DIELECTRIC ENVIRONMENT

Figure C6 shows the ∼204 nm excitation UVRR spectra of butyramide (Sigma Aldrich, ≥98% purity) in water and acetonitrile (Acros Organics, HPLC, far-UV grade) in the region of 950 cm^{-1} to 1250 cm^{-1}. In the water spectrum (Figure C6a), there are prominent bands at ∼1132 cm^{-1} and ∼1076 cm^{-1}, as well as a shoulder feature at ∼1050 cm^{-1}. Based on our DFT calculations (see Figure C6), we assign the band at 1132 cm^{-1} to a C–C stretching mode that contains significant CH_3 and NH_2 rocking character. The ∼1076 cm^{-1} band is assigned to the $AmIII^P$ vibration. The shoulder feature at ∼1050 cm^{-1} is assigned to a C–CH_3 stretching vibration.

As shown in Figure C6b, all the bands in the 950 cm^{-1} to 1250 cm^{-1} region downshift in acetonitrile. The $AmIII^P$ band downshifts 12 cm^{-1} to ∼1064 cm^{-1}. This is similar to that in propanamide [146], where the $AmIII^P$ band downshifts ∼18 cm^{-1} in acetonitrile compared to water. Therefore we conclude that on average the $AmIII^P$ band is downshifted ∼15 cm^{-1}

in acetonitrile compared to water.

C.5 CIRCULAR DICHROISM (CD) OF THE GLN_3 PEPTIDE

The temperature dependent CD spectra of Gln_3 are shown in Figure C7. The spectra were measured at a concentration of 0.5 mg·mL^{-1} using a Jasco-715 spectropolarimeter with a 0.1 cm pathlength cuvette. We averaged six individual CD spectra for each temperature.

Table C1: Summary of Crystallographic Data of Gln and Derivatives. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

	L-glutamine	D-glutamine
Chemical Formula	$C_5H_{10}N_2O_3$	$C_5H_{10}N_2O_3$
Temperature	230(2) K	230(2) K
Crystal System	orthorhombic	orthorhombic
Space Group	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Unit Cell Dimensions	a = 5.1054(4) Å, $\alpha = 90^\circ$	a = 5.1030(4) Å, $\alpha = 90^\circ$
	b = 7.7641(5) Å, $\beta = 90^\circ$	b = 7.7634(7) Å, $\beta = 90^\circ$
	c = 15.9993(11) Å, $\gamma = 90^\circ$	c = 16.0056(13) Å, $\gamma = 90^\circ$
Volume	634.19(8) $Å^3$	634.09(9) $Å^3$
Z	4	4
	N-Acetyl-L-glutamine	L-glycyl-L-glutamine
Chemical Formula	$C_7H_{13}N_2O_4$	$C_7H_{15}N_3O_5$
Temperature	230(2) K	230(2) K
Crystal System	orthorhombic	orthorhombic
Space Group	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Unit Cell Dimensions	a = 5.0732(47) Å, $\alpha = 90^\circ$	a = 5.4025(2) Å, $\alpha = 90^\circ$
	b = 12.9033(95) Å, $\beta = 90^\circ$	b = 11.5771(4) Å, $\beta = 90^\circ$
	c = 13.7830(104) Å, $\gamma = 90^\circ$	c = 15.4651(5) Å, $\gamma = 90^\circ$
Volume	902.274(2)	967.27(6) $Å^3$
Z	4	4
	L-seryl-L-asparagine	
Chemical Formula	$C_7H_{15}N_3O_6$	
Temperature	230(2) K	
Crystal System	triclinic	
Space Group	P 1	
Unit Cell Dimensions	a = 4.7643(8) Å, $\alpha = 63.988(5)^\circ$	
	b = 7.5452(12) Å, $\beta = 76.646(5)^\circ$	
	c = 8.5942(14) Å, $\gamma = 75.190(7)^\circ$	
Volume	265.90(8) $Å^3$	
Z	1	

Table C2: Summary of Crystallographic Data for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

Chemical Formula	$C_9H_{18}N_2O_3 \cdot HCl$
Temperature	230(2) K
Wavelength	1.54178 Å
Crystal Size	0.04 mm × 0.140 mm × 0.160 mm
Crystal Habit	clear, colorless rectangular prism
Crystal System	monoclinic
Space Group	P 1 2_1 1
Unit Cell Dimensions	a = 10.4579(4) Å, $\alpha = 90^\circ$
	b = 5.2517(2) Å, $\beta = 90.1490(10)^\circ$
	c = 11.6493(4) Å, $\gamma = 90^\circ$
Volume	639.80(4) $Å^3$
Z	2
Density (calc.)	1.239 $g \cdot cm^3$
Absorption Coefficient	2.601 mm^{-1}
F(000)	256

Table C3: Data Collection and Structure Refinement for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

Theta range (°)	3.79°-68.25°
Index ranges	-12≤h≤12, -6≤k≤6, -14≤l≤14
Reflections collected	12444
Independent reflections	2246 [R_{int} = 0.0532]
Refinement method	Full-matrix least-squares on F^2
Refinement program	SHELXL-2014/7 [220]
Function minimized	$\sum w(F_o^2 - F_c^2)^2$
Data / restraints / parameters	2246 / 1 / 161
Goodness-of-fit on F^2	1.812
Δ/σ_{max}	0.001
Final R indices	2181 data; I>2$\sigma(I)$, R1 = 0.0609, wR2 = 0.1566
	all data, R1 = 0.0618, wR2 = 0.1575
Weighting scheme	$w = 1/[\sigma^2(F_o^2) + (0.0680P)^2]$
	where, $P = (F_o^2 + 2F_c^2)/3$
Absolute structure parameter	0.2(0)
Extinction coefficient	0.0310(50)
Largest diff. peak and hole	1.207 and -0.585 e·Å^{-3}
R.M.S. deviation from mean	0.127 e·Å^{-3}

Table C4: Atomic Coordinates and Equivalent Isotropic Atomic Displacement Parameters ($Å^2$) for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

	x/a	y/b	z/c	U(eq)[a]
Cl1	0.92521(7)	0.46852(17)	0.35473(6)	0.0280(3)
O1	0.8535(4)	0.1912(7)	0.0330(3)	0.0450(10)
N1	0.8493(4)	0.6054(10)	0.0824(3)	0.0373(9)
C1	0.8498(4)	0.4164(10)	0.0059(3)	0.0299(10)
O2	0.7073(3)	0.7877(7)	0.6433(3)	0.0400(8)
N2	0.9045(3)	0.4597(9)	0.6235(2)	0.0241(7)
C2	0.8456(4)	0.4992(10)	0.8807(3)	0.0311(9)
O3	0.5703(3)	0.4777(7)	0.6963(3)	0.0343(7)
C3	0.8066(4)	0.2834(8)	0.8016(3)	0.0239(8)
C4	0.7839(4)	0.3639(8)	0.6766(3)	0.0215(8)
C5	0.6835(4)	0.5721(8)	0.6684(3)	0.0227(8)
C6	0.4556(4)	0.6448(10)	0.7142(5)	0.0379(10)
C7	0.4806(6)	0.8073(17)	0.8189(6)	0.0632(18)
C8	0.3495(6)	0.4559(16)	0.7351(12)	0.107(4)
C9	0.4293(7)	0.8023(19)	0.6099(6)	0.067(2)

[a]U(eq) is defined as one third of the trace of the orthogonalized U_{ij} tensor.

Table C5: Anisotropic Atomic Displacement Parameters[a] ($Å^2$) for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
Cl1	0.0312(5)	0.0323(5)	0.0204(5)	0.0026(4)	-0.0058(3)	-0.0052(4)
O1	0.070(3)	0.038(2)	0.0274(15)	0.0104(15)	-0.0012(15)	0.0098(17)
N1	0.051(2)	0.041(3)	0.0205(18)	-0.0004(18)	-0.0050(15)	-0.001(2)
C1	0.0295(18)	0.041(3)	0.0196(18)	0.0031(16)	-0.0031(14)	0.0021(17)
O2	0.0344(16)	0.0229(15)	0.063(2)	0.0082(15)	0.0047(15)	0.0011(14)
N2	0.0280(15)	0.0262(16)	0.0181(14)	0.0022(15)	-0.0003(11)	0.0010(16)
C2	0.044(2)	0.031(2)	0.0182(16)	0.0027(17)	-0.0071(15)	-0.003(2)
O3	0.0249(13)	0.0205(12)	0.0576(18)	0.0017(17)	0.0039(11)	0.0021(14)
C3	0.0297(18)	0.0231(17)	0.0189(16)	0.0056(15)	-0.0012(14)	0.0006(15)
C4	0.0239(19)	0.0230(18)	0.0175(16)	-0.0002(14)	-0.0045(14)	-0.0017(14)
C5	0.0264(18)	0.0189(17)	0.0227(16)	-0.0006(14)	-0.0051(14)	-0.0013(14)
C6	0.0252(19)	0.027(2)	0.062(3)	0.002(2)	0.0069(18)	0.0020(18)
C7	0.059(3)	0.072(4)	0.059(3)	-0.015(3)	0.007(3)	0.022(3)
C8	0.036(3)	0.035(3)	0.251(13)	0.010(6)	0.048(5)	0.000(3)
C9	0.050(3)	0.090(5)	0.060(4)	0.004(4)	-0.009(3)	0.039(4)

[a]The anisotropic atomic displacement factor exponent takes the form: $-2\pi^2[h^2 \cdot a^2 \cdot U_{11} + \cdots + 2h \cdot k \cdot a \cdot b \cdot U_{12}]$

Table C6: Hydrogen Atomic Coordinates and Isotropic Atomic Displacement Parameters ($Å^2$) for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

	x/a	y/b	z/c	U(eq)
H1NB	0.858(5)	0.744(15)	1.062(5)	0.029(16)
H1NA	0.870(7)	0.582(15)	1.156(7)	0.053(18)
H2NC	0.889(4)	0.477(13)	0.546(4)	0.024(10)
H2NB	0.926(6)	0.597(14)	0.656(5)	0.032(14)
H2NA	0.980(6)	0.355(12)	0.645(5)	0.027(13)
H2A	0.7846	0.6398	0.8723	0.037
H2B	0.9301	0.5618	0.8581	0.037
H3A	0.7281	0.2061	0.8312	0.029
H3B	0.8736	0.1529	0.8033	0.029
H7A	0.4972	0.6984	0.8845	0.095
H7B	0.5542	0.9153	0.8053	0.095
H7C	0.4063	0.9126	0.8342	0.095
H8A	0.369	0.3564	0.8031	0.161
H8B	0.2696	0.5463	0.7463	0.161
H8C	0.3418	0.3434	0.6694	0.161
H9A	0.355	0.9086	0.6234	0.1
H9B	0.5027	0.9092	0.5938	0.1
H9C	0.4131	0.6913	0.545	0.1
H4	0.749(5)	0.212(10)	0.627(4)	0.020(12)

Table C7: Bond Lengths (Å) and Bond Angles (°) for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

Bond Lengths			
O1-C1	1.225(6)	N1-C1	1.334(7)
N1-H1NB	0.77(8)	N1-H1NA	0.89(8)
C1-C2	1.522(5)	O2-C5	1.195(6)
N2-C4	1.494(5)	N2-H2NC	0.93(5)
N2-H2NB	0.85(7)	N2-H2NA	1.00(6)
C2-C3	1.515(6)	C2-H2A	0.98
C2-H2B	0.98	O3-C5	1.325(5)
O3-C6	1.501(5)	C3-C4	1.534(5)
C3-H3A	0.98	C3-H3B	0.98
C4-C5	1.519(5)	C4-H4	1.05(5)
C6-C9	1.494(9)	C6-C8	1.509(8)
C6-C7	1.511(9)	C7-H7A	0.97
C7-H7B	0.97	C7-H7C	0.97
C8-H8A	0.97	C8-H8B	0.97
C8-H8C	0.97	C9-H9A	0.97
C9-H9B	0.97	C9-H9C	0.97
Bond Angles			
C1-N1-H1NB	119.(5)	C1-N1-H1NA	123.(5)
H1NB-N1-H1NA	114.(7)	O1-C1-N1	123.1(4)
O1-C1-C2	121.6(4)	N1-C1-C2	115.3(4)
C4-N2-H2NC	107.(3)	C4-N2-H2NB	109.(4)
H2NC-N2-H2NB	114.(6)	C4-N2-H2NA	113.(3)
H2NC-N2-H2NA	115.(5)	H2NB-N2-H2NA	98.(5)
C3-C2-C1	112.1(4)	C3-C2-H2A	109.2
C1-C2-H2A	109.2	C3-C2-H2B	109.2
C1-C2-H2B	109.2	H2A-C2-H2B	107.9
C5-O3-C6	122.0(4)	C2-C3-C4	114.3(3)
C2-C3-H3A	108.7	C4-C3-H3A	108.7
C2-C3-H3B	108.7	C4-C3-H3B	108.7
H3A-C3-H3B	107.6	N2-C4-C5	108.4(3)
N2-C4-C3	110.9(3)	C5-C4-C3	111.3(3)
N2-C4-H4	109.(3)	C5-C4-H4	106.(3)
C3-C4-H4	111.(3)	O2-C5-O3	126.9(4)
O2-C5-C4	123.6(4)	O3-C5-C4	109.4(3)
C9-C6-O3	110.8(4)	C9-C6-C8	111.2(7)
O3-C6-C8	103.1(4)	C9-C6-C7	112.0(6)
O3-C6-C7	107.8(4)	C8-C6-C7	111.5(7)
C6-C7-H7A	109.5	C6-C7-H7B	109.5
H7A-C7-H7B	109.5	C6-C7-H7C	109.5
H7A-C7-H7C	109.5	H7B-C7-H7C	109.5
C6-C8-H8A	109.5	C6-C8-H8B	109.5
H8A-C8-H8B	109.5	C6-C8-H8C	109.5
H8A-C8-H8C	109.5	H8B-C8-H8C	109.5
C6-C9-H9A	109.5	C6-C9-H9B	109.5
H9A-C9-H9B	109.5	C6-C9-H9C	109.5
H9A-C9-H9C	109.5	H9B-C9-H9C	109.5

Table C8: Raman Frequencies (cm^{-1}) and Assignments of Crystalline L-glutamine[a]. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

Expt.	Calc.	PED[b] (≥5% contribution)	Assignment
1205	1215	$-\tau C_\beta H_2$ (21), $-\nu C_\alpha C_\beta$ (18), $-\tau C_\gamma H_2$ (16), $-\rho' NH_3$ (13), $\delta NC_\alpha C(OO)$ (5)	$\tau CH_2/\nu C$–C
1166	1153	$-\rho' C_\alpha H$ (20), $\tau C_\gamma H_2$ (17), $-\rho' NH_3$ (13), $\tau C_\beta H_2$ (12), $-\nu C_\alpha C_\beta$ (8), $\rho C_\gamma H_2$ (6)	$\rho CH/\tau CH_2$
1135	1122	$\nu C_\beta C_\gamma$ (34), $\rho N_\epsilon H_2$ (17), $-\nu C_\alpha C_\beta$ (7), $\nu C_\alpha N$ (6), $-\beta N_\epsilon C_\delta C_\gamma$ (5)	νC–C/ρNH_2
1105	1109	ρNH_3 (27), $-\rho' C_\alpha H$ (10), $-\rho N_\epsilon H_2$ (10), $-\delta' NC_\alpha C(OO)$ (9), $-\rho C_\alpha H$ (7), $\nu C_\alpha N$ (7)	ρNH_3
1097	1097	$\nu C_\beta C_\gamma$ (26), $-\rho N_\epsilon H_2$ (26), $-\nu C_\delta N_\epsilon$ (13), $-\rho NH_3$ (8)	AmIII[P]
1086	1038	$\nu C_\alpha N$ (36), $-\nu C_\beta C_\gamma$ (9), $\rho C_\beta H_2$ (8), $\rho' NH_3$ (6), $\rho C_\gamma H_2$ (5), $\rho' C_\alpha H$ (5)	νC–N
1052	1003	ρNH_3 (25), $-\nu C_\alpha N$ (19), $\rho C_\gamma H_2$ (14), $\rho C_\beta H_2$ (14), $\nu C_\alpha C_\beta$ (7), $-\tau C_\beta H_2$ (5)	$\rho NH_3/\nu$C–N
1000	974	$-\rho' NH_3$ (38), $\nu C_\alpha C_\beta$ (25), $-\nu C_\alpha C$ (8), $-\sigma CC_\alpha C_\beta$ (7), $\nu C_\alpha N$ (6)	$\rho NH_3/\nu$C–C

[a]Frequencies correspond to visible Raman (633 nm excitation) spectrum. [b]ν: stretch; δ_s: symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

Table C9: Raman Frequencies (cm^{-1}) and Assignments of Crystalline D-glutamine[a]. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

Expt.	Calc.	PED[b] (≥5% contribution)	Assignment
1204	1215	$-\tau C_\beta H_2$ (21), $-\nu C_\alpha C_\beta$ (18), $-\tau C_\gamma H_2$ (16), $-\rho' NH_3$ (13), $\delta NC_\alpha C(OO)$ (5)	$\tau CH_2/\nu$C–C
1165	1153	$-\rho' C_\alpha H$ (20), $\tau C_\gamma H_2$ (17), $-\rho' NH_3$ (13), $\tau C_\beta H_2$ (12), $-\nu C_\alpha C_\beta$ (8), $\rho C_\gamma H_2$ (6)	$\rho CH/\tau CH_2$
1134	1122	$\nu C_\beta C_\gamma$ (34), $\rho N_\epsilon H_2$ (17), $-\nu C_\alpha C_\beta$ (7), $\nu C_\alpha N$ (6), $-\beta N_\epsilon C_\delta C_\gamma$ (5)	νC–C/ρNH_2
1105	1109	ρNH_3 (27), $-\rho' C_\alpha H$ (10), $-\rho N_\epsilon H_2$ (10), $-\delta' NC_\alpha C(OO)$ (9), $-\rho C_\alpha H$ (7), $\nu C_\alpha N$ (7)	ρNH_3
1096	1097	$\nu C_\beta C_\gamma$ (26), $-\rho N_\epsilon H_2$ (26), $-\nu C_\delta N_\epsilon$ (13), $-\rho NH_3$ (8)	AmIII[P]
1086	1038	$\nu C_\alpha N$ (36), $-\nu C_\beta C_\gamma$ (9), $\rho C_\beta H_2$ (8), $\rho' NH_3$ (6), $\rho C_\gamma H_2$ (5), $\rho' C_\alpha H$ (5)	νC–N
1051	1003	ρNH_3 (25), $-\nu C_\alpha N$ (19), $\rho C_\gamma H_2$ (14), $\rho C_\beta H_2$ (14), $\nu C_\alpha C_\beta$ (7), $-\tau C_\beta H_2$ (5)	$\rho NH_3/\nu$C–N
999	974	$-\rho' NH_3$ (38), $\nu C_\alpha C_\beta$ (25), $-\nu C_\alpha C$ (8), $-\sigma CC_\alpha C_\beta$ (7), $\nu C_\alpha N$ (6)	$\rho NH_3/\nu$C–C

[a]Frequencies correspond to visible Raman (633 nm excitation) spectrum. [b]ν: stretch; δ_s: symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

Table C10: Raman Frequencies (cm^{-1}) and Assignments of Crystalline N-Acetyl-L-glutamine[a]. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

Expt.	Calc.	PED[b] (≥5% contribution)	Assignment
	1192	$-\tau C_\beta H_2$ (32), $-\tau C_\gamma H_2$ (24), $-\nu C_\alpha C_\beta$ (12), $\rho' C_\alpha H$ (11)	τCH_2
1180	1166	$\nu C_\alpha N$ (46), $-\delta NC_\alpha C_\beta COO$ (8), $\omega C_\beta H_2$ (7), $-\sigma C_\alpha C_\beta C_\gamma$ (6), $\nu C_\beta C_\gamma$ (5)	νC–N
1138	1119	$\nu C_\alpha C_\beta$ (22), $-\nu C_\alpha N$ (12), $\nu C_\beta C_\gamma$ (10), $\rho N_\epsilon H_2$ (8), $-\tau C_\gamma H_2$ (5)	νC–C/νC–N
1111	1110	$\rho N_\epsilon H_2$ (20), $-\nu C_\alpha C_\beta$ (15), $\nu C_\beta C_\gamma$ (14), $\rho C_\beta H_2$ (7), $\rho C_\gamma H_2$ (6), $\nu C_\delta N_\epsilon$ (5), $-\omega C_\gamma H_2$ (5)	$\rho NH_2/\nu$C–C
1071	1093	$-\nu C_\beta C_\gamma$ (42), $\rho N_\epsilon H_2$ (24), $\nu C_\delta N_\epsilon$ (13)	AmIII[P]
1061	1058	ρCH_3 (63), $-\delta_s NC'O'CH_3$ (19), $-\delta_{as} CH_3$ (9), $-\rho' CH_3$ (7)	ρCH_3
1022	1043	$\rho C_\beta H_2$ (25), $\rho' C_\alpha H$ (15), $\rho C_\gamma H_2$ (14), $-\nu C_\alpha C$ (8), $\nu C_\alpha C_\beta$ (8), $\delta CC_\alpha C_\beta$ (5)	$\rho CH_2/\rho CH$
	1013	$\rho' CH_3$ (58), $-\nu C'CH_3$ (12), ρCH_3 (5), $-\delta_{as}' CH_3$ (5)	ρCH_3
997	971	$\nu C'CH_3$ (28), $\nu C'N$ (16), $\rho C'O$ (11), $-\nu C_\alpha C_\beta$ (10), $-\beta C_\alpha NC'$ (8)	νC–CH_3/νC–N

[a]Frequencies correspond to visible Raman (633 nm excitation) spectrum. [b]ν: stretch; δ_s: symmetric deformation; δ_{as}: asymmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

Table C11: Raman Frequencies (cm^{-1}) and Assignments of Crystalline L-glutamine t-butyl ester HCl[a]. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

Expt.	Calc.	PED[b] (≥5% contribution)	Assignment
1195	1195	νCO (17), ρCH_3 (10), $-\rho CH_3$ (10), $-\rho CH_3$ (9), $-\nu$CO (6), $-\delta$C'O'CCC (6), δ''C'O'CCC (6), $-\delta$'C'O'CCC (6), ρCH_3 (5), ρCH_3 (5)	$\nu CO/\rho CH_3$
1151	1161	$\rho C_\alpha H$ (23), ρNH_3 (19), $\tau C_\gamma H_2$ (11), $-\nu C_\alpha C_\beta$ (10), $\tau C_\beta H_2$ (8)	$\rho CH/\rho NH_3$
1117	1126	$-\rho NH_3$ (19), $-\nu C_\beta C_\gamma$ (16), $-\nu C_\alpha N$ (14), $\delta NC_\alpha C_\beta C$ (9), $-\rho N_\epsilon H_2$ (5)	$\rho NH_3/\nu C–C$
1105	1119	$-\rho NH_3$ (15), $\rho N_\epsilon H_2$ (14), $\rho C_\beta H_2$ (10), $\nu C_\beta C_\gamma$ (9), $\rho C_\gamma H_2$ (7), $\rho C_\alpha H$ (6), $\delta NC_\alpha C_\beta C$ (6)	$\rho NH_3/\rho NH_2$
1082	1100	$\rho N_\epsilon H_2$ (34), $-\nu C_\beta C_\gamma$ (24), $\nu C_\delta N_\epsilon$ (14), $\nu C_\delta O_\epsilon$ (5)	AmIII[P]
1043	1068	$-\nu C_\alpha N$ (17), ρCH_3 (13), $-\rho CH_3$ (11), $-\rho CH_3$ (11), $\nu C_\beta C_\gamma$ (10), ρCH_3 (5), $\nu C_\alpha C$ (5)	$\nu C–N/\rho CH_3$
1030	1058	ρCH_3 (27), $-\rho CH_3$ (26), $-\rho CH_3$ (10), ρCH_3 (10), $-\nu CCH_3$ (6)	ρCH_3
	1057	$-\nu C_\alpha N$ (24), $\nu C_\beta C_\gamma$ (11), ρCH_3 (9), ρCH_3 (8), $-\rho CH_3$ (7), $-\rho' CH_3$ (5)	$\nu C–N/\nu C–C$
	1014	$-\rho NH_3$ (24), $-\rho C_\beta H_2$ (18), $-\rho C_\gamma H_2$ (16), $\nu C_\alpha N$ (14), $-\nu C_\alpha C_\beta$ (5), $\tau C_\beta H_2$ (5), $-\delta_s N_\epsilon C_\delta(O_\epsilon)C_\gamma$ (5)	$\rho NH_3/\rho CH_2$
998	989	ρNH_3 (35), $\nu C_\alpha C_\beta$ (25), $-\nu C_\alpha C$ (11)	$\rho NH_3/\nu C–C$

[a]Frequencies correspond to visible Raman (633 nm excitation) spectrum. [b]ν: stretch; δ_s: symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

Table C12: Raman Frequencies (cm^{-1}) and Assignments of Crystalline L-glycyl-L-glutamine[a]. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

Expt.	Calc.	PED[b] (≥5% contribution)	Assignment
	1201	τ'$C_\beta H_2$ (33), $-\rho$'$C_\alpha H$ (24), τ'$C_\gamma H_2$ (15), $\tau C_\gamma H_2$ (6), $\nu C_\alpha C_\beta$ (5)	$\tau CH_2/\rho CH$
1171	1147	$-\nu C_\alpha N$ (37), $\delta NC_\alpha C$ (7), $-\rho N'H_3$ (7), $-\nu C_{\alpha'} N'$ (7)	$\nu C–N$
1139	1138	$\nu C_\alpha C_\beta$ (26), $-\nu C_\beta C_\gamma$ (19), $-\rho C_\beta H_2$ (8), $-\rho N_\epsilon H_2$ (6), $-\delta C_\alpha C_\beta C$ (6), $-\nu C_\alpha N$ (6), $-\omega C_\beta H_2$ (5)	$\nu C–C$
1124	1128	$\rho' N'H_3$ (47), $\tau' C_{\alpha'} H_2$ (18), $-\rho C_{\alpha'} H_2$ (13), $-\delta_s C'NC_\alpha C(OO)$ (5)	ρNH_3
	1105	$\rho N'H_3$ (29), $\rho N_\epsilon H_2$ (12), $-\tau C_{\alpha'} H_2$ (7), $-\tau' C_{\alpha'} H_2$ (6), $\nu C_{\alpha'} N'$ (6), $-\sigma N'C_{\alpha'} C'$ (5), $-\nu C_\alpha N$ (5)	ρNH_3
1093	1100	$\rho N_\epsilon H_2$ (33), $\nu C_\delta N_\epsilon$ (13), $-\rho N'H_3$ (12), $-\nu C_\beta C_\gamma$ (6), $\nu C_\delta O_\epsilon$ (5)	AmIII[P]
1082	1081	$-\nu C_{\alpha'} N'$ (14), $\nu C_\beta C_\gamma$ (13), $\nu C' C_{\alpha'}$ (8), $\rho C_\gamma H_2$ (7), $\nu C_\alpha C_\beta$ (7), $\rho' C_\alpha H$ (6), $-\nu C_\alpha N$ (5)	$\nu C–N/\nu C–C$
1048	1045	$-\nu C_{\alpha'} N'$ (56), $\nu C_\alpha N$ (9), $-\nu C_\alpha C_\beta$ (7), $-\nu C_\beta C_\gamma$ (5)	$\nu C–N$
1027	1013	$\rho C_\beta H_2$ (28), $\delta C_\alpha C_\beta C$ (9), $\tau C_\gamma H_2$ (8), $-\nu C_\beta C_\gamma$ (7), $\rho C_\gamma H_2$ (5), $\delta NC_\alpha C'$ (5)	ρCH_2
997	946	$\nu C' C_{\alpha'}$ (19), $-\rho C_\gamma H_2$ (10), $\rho C' O'$ (9), $-\beta NC' C_\alpha$ (6), $-\beta C' C_{\alpha'} N$ (6), $-\rho N'H_3$ (5)	$\nu C–C/\rho CH_2$

[a]Frequencies correspond to visible Raman (633 nm excitation) spectrum. [b]ν: stretch; δ_s: symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

Table C13: Raman Frequencies (cm^{-1}) and Assignments of Crystalline L-seryl-L-asparagine[a]. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

Expt.	Calc.	PED[b] (≥5% contribution)	Assignment
1188	1184	$-\rho C_{\beta'}H_2$ (26), $\rho^{'}N^{'}H_3$ (14), $-\delta NC_{\alpha'}C'$ (9), $-\tau C_{\beta'}H_2$ (8), $\beta C_{\beta'}O_{\gamma'}H$ (7), $\rho^{'}C_{\alpha'}H$ (6), $\nu C_{\alpha'}N'$ (5), $\rho N^{'}H_3$ (5)	$\rho CH_2/\rho NH_3$
1159	1164	$-\nu C_\alpha N$ (37), $\rho^{'}N^{'}H_3$ (7), $\nu C_\alpha C_\beta$ (7), $\delta CC_\alpha N$ (5), $\rho^{'}C_\alpha H$ (5)	νC–N
1121	1119	$-\nu C_{\beta'}O_{\gamma'}H$ (17), $\rho N_\delta H_2$ (13), $\nu C_\alpha C_\beta$ (13), $\nu C_{\alpha'}N'$ (8), $-\rho^{'}C_{\alpha'}H$ (6), $\rho N^{'}H_3$ (5), $-\rho^{'}N^{'}H_3$ (5)	νC–OH/ρNH_2
1108	1116	$\nu C_{\beta'}O_{\gamma'}H$ (66), $\rho N_\delta H_2$ (9)	νC–OH
1083	1107	$-\rho N_\delta H_2$ (16), $\nu C_{\alpha'}N'$ (10), $-\rho^{'}C_{\alpha'}H$ (10), $\nu C_{\beta'}O_{\gamma'}H$ (8), $\rho N^{'}H_3$ (8), $-\nu C_\alpha N$ (6), $-\nu C_\delta N_\delta$ (5)	$\rho NH_2/\nu$C–N
1065	1083	$-\rho N^{'}H_3$ (20), $\nu C_{\alpha'}N'$ (10), $-\nu C^{'}C_{\alpha'}$ (9), $\nu C_\alpha C_\beta$ (8), $\rho C_{\alpha'}H$ (7), $-\nu C_\beta O_{\gamma'}H$ (5), $\rho^{'}C_{\alpha'}H$ (5)	$\rho NH_3/\nu$C–N
1051	1069	$\nu C_\alpha C_\beta$ (23), $-\nu C_{\alpha'}N'$ (11), $\nu C_\alpha N$ (9), $-\delta CC_\alpha N$ (7), $-\sigma C_\alpha C_\beta C_\gamma$ (5), $-\rho N_\delta H_2$ (5)	$AmIII^P$
1019	1023	$-\rho C_\beta H_2$ (23), $\delta C_\beta C_\alpha C$ (12), $\delta C^{'}C_\alpha N$ (8), $-\nu C_\alpha C$ (7), $\delta_s C_\alpha COO$ (6), $\beta C_\alpha NC'$ (6), $\rho^{'}C_\alpha H$ (5)	$\rho CH_2/\delta CCC$
1004	981	$\rho C_{\beta'}H_2$ (28), $\rho^{'}N^{'}H_3$ (25), $-\nu C_{\alpha'}C_{\beta'}$ (14), $\rho N^{'}H_3$ (10)	$\rho CH_2/\rho NH_3$

[a]Frequencies correspond to visible Raman (633 nm excitation) spectrum. [b]ν: stretch; δ_s: symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

Table C14: UVRR Frequencies (cm^{-1}) and Assignments of Butyramide in Acetonitrile and Water. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

H_2O	CH_3CN	Calc.	PED[b] (≥5% contribution)	Assignment
1132	1125	1143	$\nu C_\beta C_\gamma$ (27), $\rho' C_\alpha H_3$ (19), $\rho N_\epsilon H_2$ (14), $-\sigma C_\alpha C_\beta C_\gamma$ (8), $-\sigma C_\beta C_\gamma C_\delta$ (6), $\omega C_\beta H_2$ (6) $-\beta C_\gamma C_\delta N_\epsilon$ (5)	νC–C/ρCH_3
		1126	$-\rho C_\alpha H_3$ (17), $\rho C_\beta H_2$ (17), $\rho C_\gamma H_2$ (17), $\tau C_\gamma H_2$ (16), $-\delta_s N_\epsilon C_\delta C_\gamma$ (7), $\tau C_\beta H_2$ (7) $-\nu C_\beta C_\gamma$ (6)	$\rho CH_3/\rho CH_2$
1076	1064	1092	$\rho N_\epsilon H_2$ (37), $\nu C_\delta N_\epsilon$ (17), $-\nu C_\beta C_\gamma$ (14), $-\tau C_\gamma H_2$ (6)	$AmIII^P$
1050	1040	1067	$\nu C_\alpha C_\beta$ (59), $-\nu C_\beta C_\gamma$ (10), $\rho' C_\alpha H_3$ (9), $-\rho N_\epsilon H_2$ (6), $\omega C_\gamma H_2$ (6)	νC–CH_3

[a]Frequencies correspond to UVRR (204 nm excitation) spectra. [b]ν: stretch; δ_s: symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

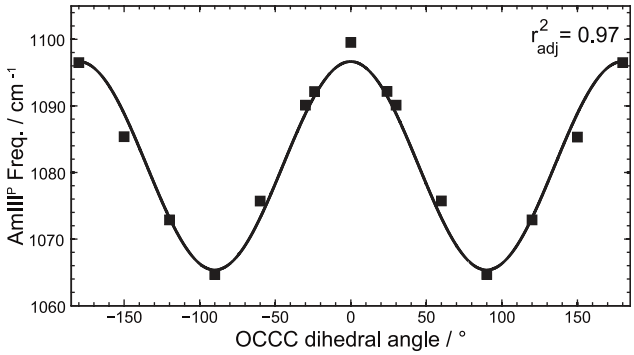


Figure C1: AmIII^P frequency dependence on the OCCC dihedral angle of butyramide from DFT calculations. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

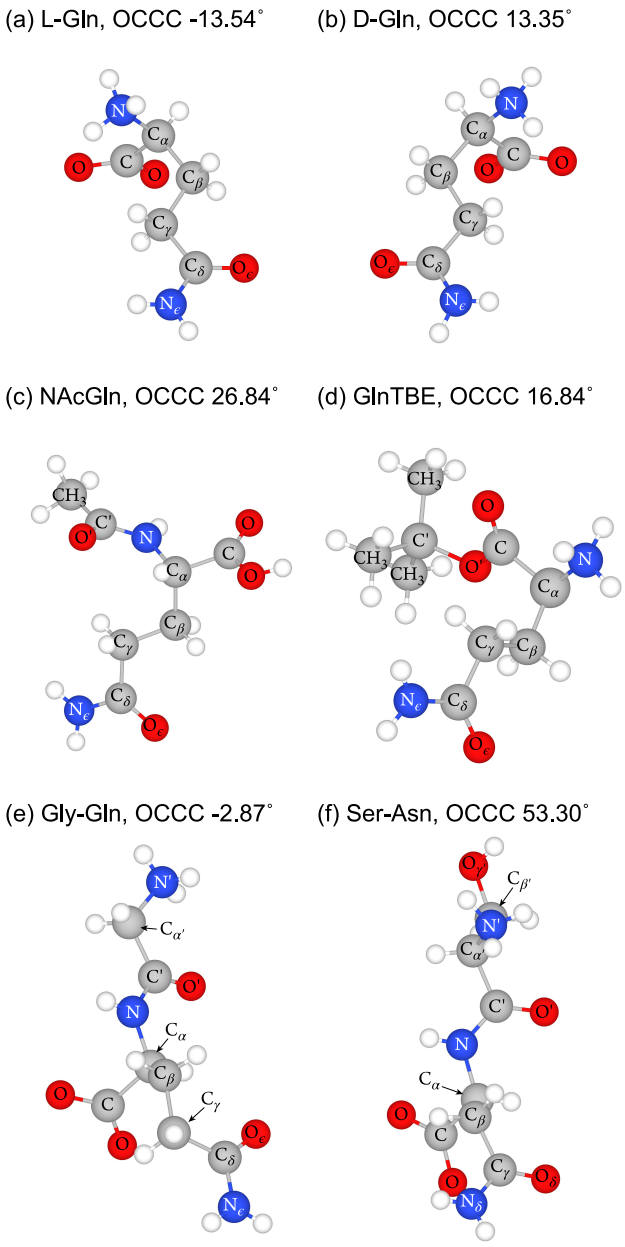


Figure C2: Crystal structures of Gln and Gln derivatives examined in this study. The atomic labeling schemes shown are used for Table C8–Table C13. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

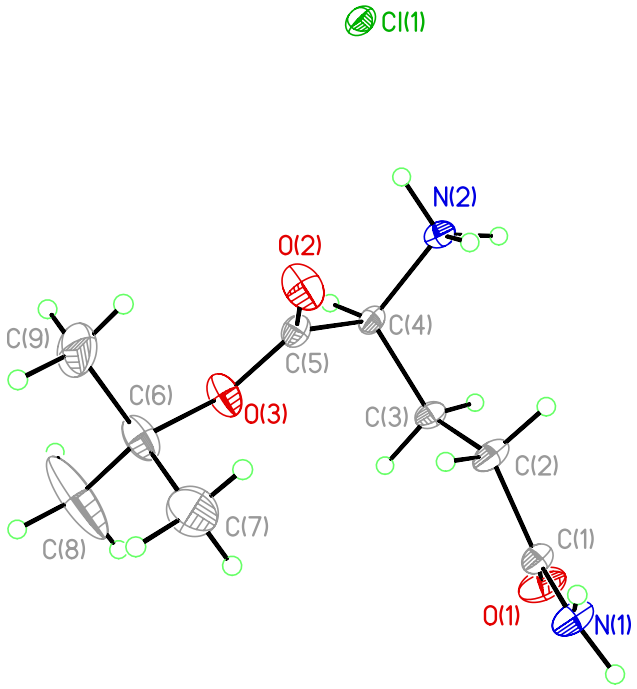


Figure C3: OTREP diagram of GlnTBE with atomic labeling scheme used in Table C4–Table C7. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

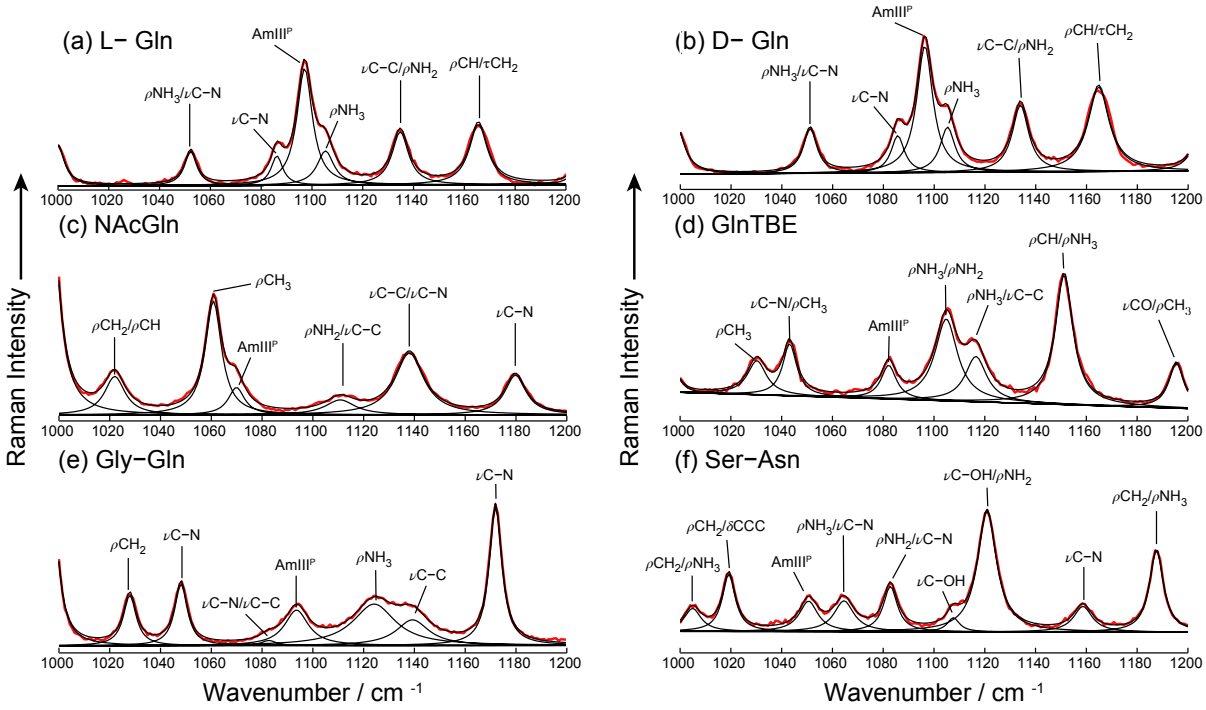


Figure C4: Visible Raman (633 nm excitation) spectra of compounds examined in this study. (a) L-Gln; (b) D-Gln; (c) NAcGln; (d) GlnTBE; (e) Gly-Gln; and (f) Ser-Asn. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

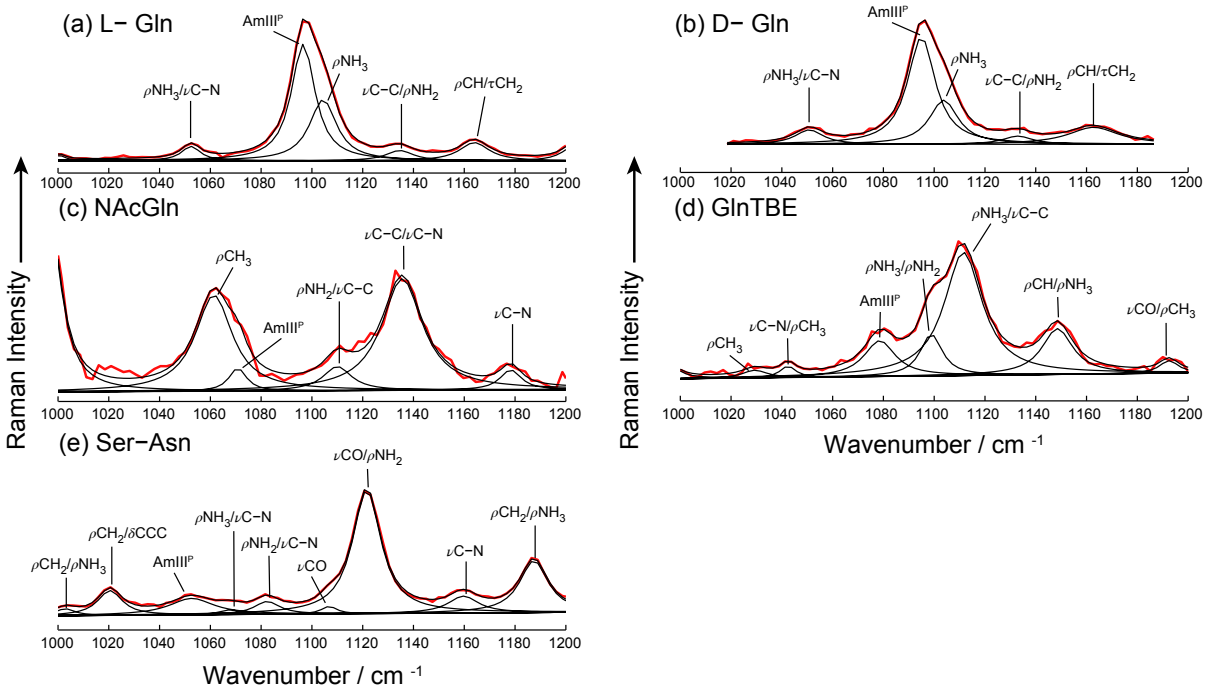


Figure C5: UVRR (229 nm excitation) spectra of compounds examined in this study. (a) L-Gln; (b) D-Gln; (c) NAcGln; (d) GlnTBE; and (e) Ser-Asn. CAdapted with permission from [180]. Copyright © (2015), American Chemical Society.

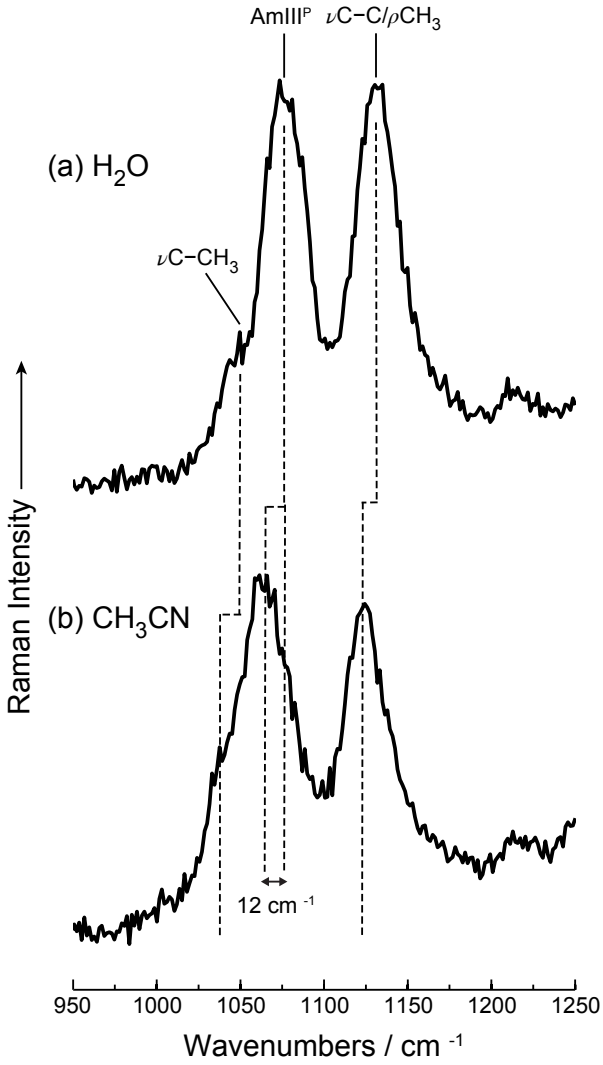


Figure C6: UVRR (204 nm excitation) spectra of buytramide in (a) water and (b) acetonitrile. The AmIII$^{\mathrm{P}}$ of butyramide downshifts 12 cm^{-1} as the solvent transfers from acetonitrile to water. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

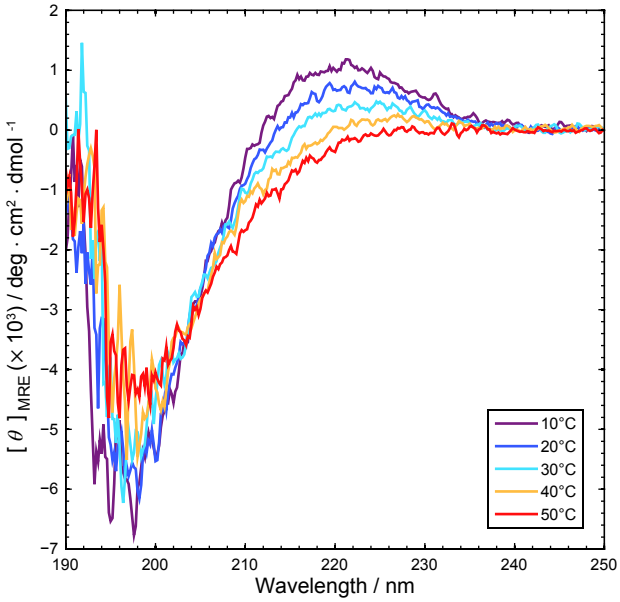


Figure C7: Temperature dependent Circular Dichroism (CD) spectra of Q_3. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

APPENDIX D

SUPPORTING INFORMATION FOR CHAPTER 5.0

D.1 EXPERIMENTAL SECTION

D.1.1 UVRR spectroscopy of fibril films

Solutions of NDQ10 and DQ10 fibril aggregates were aliquoted onto the grooves of brass cylindrical cells. The solutions were dried for ∼5 h, after which a film could be observed. The UVRR instrumentation used to collect spectra of NDQ10 and DQ10 fibril films was the same as described in the main text. The laser light was focused onto the grooves of the brass cells, which were spun in order to prevent accumulation of photochemical or thermal degradation products. UVRR spectra were collected using both ∼197 nm and ∼204 nm excitation.

D.1.2 UVRR Spectral Processing

All UVRR spectra were processed using home-written MATLAB scripts in order to remove cosmic rays, average and calibrate spectra, as well as subtract the spectral contributions of water (*e.g.* H_2O, D_2O, and HDO) and Suprasil quartz from NMR tubes. The spectra were calibrated using the 801.3 cm^{-1}, 1028.3 cm^{-1}, 1157.6 cm^{-1}, 1266.4 cm^{-1}, and 1444.4 cm^{-1} bands of cyclohexane. The spectral contributions of water and quartz were removed using a method similar to that described by Banerjee and coworkers [222, 223].

To subtract the contributions of water and quartz, we first calculated the first-derivatives of the spectra. The relative contributions of water and quartz in the raw spectra were found

via a classical multiple linear least-squares regression such that:

$$\mathbf{S'} = \mathbf{K}\mathbf{S'_P} \tag{D.1}$$

where $\mathbf{S'}$ is the $(n \times 1)$ row vector that represents the first-derivative (denoted by $'$) of the experimentally measured raw spectrum, and $\mathbf{S'_P}$ is the $(n \times m)$ matrix composed of row vector elements that contain the first-derivative spectra of the pure water and quartz spectra. The $(m \times 1)$ row vector, $\mathbf{K}$, contains the least-squares scaling coefficients for each of the different background spectral components. The least-squares solution of $\mathbf{K}$ for eq. D.1 is [224]:

$$\mathbf{K} = (\mathbf{S'_P}^T\mathbf{S'_P})^{-1}\mathbf{S'_P}^T\mathbf{S'} \tag{D.2}$$

where $(\mathbf{S'_P}^T\mathbf{S'_P})^{-1}\mathbf{S'_P}^T$ is the pseudo-inverse matrix of $\mathbf{S'_P}$. Banerjee and coworkers [222, 223] discuss the advantages of using the first-derivative spectra to determine relative contributions of different spectral components. Using the first-derivative spectra is most advantageous for spectra that contain multiple, overlapping spectral components.

After determining $\mathbf{K}$, eq. D.3 was utilized to subtract the contributions of water and quartz to obtain the spectrum of the analyte of interest:

$$\mathbf{S_{analyte}} = \mathbf{S} - \mathbf{K}\mathbf{S_P} \tag{D.3}$$

where $\mathbf{S_{analyte}}$, $\mathbf{S}$, and $\mathbf{S_P}$ are the zeroth-derivative analyte, raw, and water/quartz UVRR spectra, respectively.

D.1.3 UVRR Spectral Peak Fitting

The GRAMS AI software suite (ver. 8.0, Thermo Fisher Scientific) was used to peak fit the UVRR spectra. The spectra, $S(\nu)$, were parsimoniously fit as the sum of pure Gaussian and Lorentzian bands. *i.e.*

$$S(\nu) = \sum_i \left[f_i H_i e^{-\left(\frac{\nu-\nu_i}{w_i}\right)^2 (4ln(2))} + (1 - f_i)\frac{H_i}{4(\frac{\nu-\nu_i}{w_i})^2 + 1} \right] \tag{D.4}$$

where $f_i = 1$ if the i^{th} band is a Gaussian, or 0 if the i^{th} band is a Lorentzian. The parameters H_i, ν_i, and w_i are the heights, center frequencies, and widths, respectively, of the i^{th} band.

D.1.4 Calculation of the Ψ and χ_3 Angle Distributions

The distributions of Ψ and χ_3 dihedral angles, shown in Figure 5.7 and Figure 5.11 in the main text, were calculated using methodologies previously described in detail [96, 97, 105, 180]. Briefly, we assume that the inhomogeneously broadened, experimentally measured AmIII_3^S and $\mathrm{AmIII}^{\mathrm{P}}$ band profiles, $B(\nu)$, can be modeled as the sum of M Lorentzian bands:

$$B(\nu) = \frac{1}{\pi} \sum_{i}^{M} \frac{p_i \Gamma^2}{\Gamma^2 + (\nu - \nu_i)^2} \tag{D.5}$$

where p_i is the probability for the i^{th} band to occur at center frequency ν_i. The band width parameter, Γ, is the homogeneous linewidth of the AmIII_3^S or $\mathrm{AmIII}^{\mathrm{P}}$ vibrations. We previously [105, 180] estimated from peptide crystals that Γ is $\sim$7.5 cm^{-1} for the $\mathrm{AmIII}_3^{\mathrm{S}}$ and $\sim$6.6 cm^{-1} $\mathrm{AmIII}^{\mathrm{P}}$. After decomposing the band profiles into Lorentzians, we then correlate the different i^{th} frequencies of the $\mathrm{AmIII}_3^{\mathrm{S}}$ and $\mathrm{AmIII}^{\mathrm{P}}$ band envelopes to their respective Ψ or χ_3 dihedral angles.

D.1.4.1 Correlating the $\mathrm{AmIII}_3^{\mathrm{S}}$ Frequencies to Ψ Angles

We used the following equation to correlate the $\mathrm{AmIII}_3^{\mathrm{S}}$ band frequencies to Ψ angles for the Figure 5.7 black distributions shown in the main text:

$$\nu_i(\Psi) = 1239\ (\mathrm{cm}^{-1}) - 54\ (\mathrm{cm}^{-1}) \sin(\Psi + 26^\circ) \tag{D.6}$$

Eq. D.6 was derived by Mikhonin *et al.* [97] for situations when there is strong peptide-peptide hydrogen bonding, such as in the case for fibril peptide bonds.

The Ψ angles for the Figure 5.7 distributions shown in blue were obtained by using:

$$\nu_i(\Psi, T) = 1250\ (\mathrm{cm}^{-1}) - 54\ (\mathrm{cm}^{-1}) \sin(\Psi + 26^\circ) + 0.06\ (\mathrm{cm}^{-1}/^\circ\mathrm{C})(T - T_0) \tag{D.7}$$

where $T = 22\,^\circ\mathrm{C}$ is the experimental temperature and $T_0 = 0\,^\circ\mathrm{C}$. Eq. D.7 was derived by Mikhonin *et al.* [97] for situations when the hydrogen bonding state of the peptide bond N–H groups is unknown. This situation occurs, for example, in the case of solvent accessible fibril peptide bonds. It is unknown if these peptide bonds are exhangeable in solvent because

they are located on the surface of fibrils, or whether they are more in more disordered regions of the aggregates.

D.1.4.2 Correlating the AmIII^{P} Frequencies to χ_3 Angles

In the case of the AmIII^{P}, we used the following equation to correlate the band frequencies to χ_3 angles for the distributions shown in Figure 5.11

$$\nu_i(\chi_3) = 1076\ (\mathrm{cm}^{-1}) + 29\ (\mathrm{cm}^{-1})\cos(2\chi_3) + 9\ (\mathrm{cm}^{-1})\cos(\chi_3 + 99^{\circ}) \tag{D.8}$$

where ν_i is the i^{th} AmIII^{P} frequency. Eq. D.8 was derived by Punihaole *et al.* [180] as an "average" expression to be used in situations when the hydrogen bonding and dielectric environments of Gln side chains are unknown. We previously showed [146] that stronger (weaker) hydrogen bonding and higher (lower) dielectric environments upshift (downshift) the AmIII^{P} frequency. In the case of Q10 fibrils, the inter-amide hydrogen bonding of the Gln side chains is strong, but the dielectric constant of the environment is also low. Thus, we utilized eq. D.8 since it averages these two competing effects. We are presently investigating which of these effects dominates the AmIII^{P} frequency in polyQ fibrils.

It should be noted that in using eq. D.8, each AmIII^{P} frequency can be correlated to as many as four possible χ_3 dihedral angles. However, as discussed in detail by Punihaole *et al.* [180], χ_3 dihedral angles that are greater than +90° and less than -90° are nearly forbidden for Gln and Asn. Thus, in using eq. D.8, we only considered the physically relevant χ_3 angles that occur between -90° and +90°.

D.2 COMPUTATIONAL SECTION

D.2.1 Density Functional Theory (DFT) Calculations

DFT calculations [135] were performed on the zwitterion form of the Gln amino acid (Figure D1) using the *GAUSSIAN 09* program [136]. The M06-2X density functional was employed using the 6-311++g** basis set [137]. Water was modeled implicitly by placing the

Gln molecule in an ellipsoidal cavity surrounded by a polarizable continuum dielectric model. The calculated frequencies were not scaled. The potential energy distribution (PED) of each vibration was obtained from the *GAUSSIAN 09* output files by employing a MATLAB script that has been previously published [180].

D.2.2 RMSD Metric

The extent of model fibril dissociation was quantified with a root mean square deviation (RMSD) metric. The equation for RMSD is shown and this metric corresponds to the spatial deviation of atoms.

$$RMSD = \frac{1}{N}\sqrt{\sum_{i=1}^{N}[(x_i - x_{ref})^2 + (y_i - y_{ref})^2 + (z_i - z_{ref})^2]} \tag{D.9}$$

where N is the number of atoms used in the RMSD calculation, x_i, y_i, z_i are the current coordinate positions of atom i, and $x_{ref}, y_{ref}, z_{ref}$ are the coordinate positions of atom i in the reference structure. Backbone atoms used were C_α, the carbonyl carbon, the carbonyl oxygen, and the peptide backbone nitrogen. For our peptide system, there are 320 atoms: four backbone atoms per residue, ten Gln residues per peptide, and eight peptides per fibril model.

Before the RMSD measurement was taken for each step, the model fibril was superimposed on the initial reference structure to eliminate the effect of fibril translation and rotation on the RMSD value. The interpretation of this metric was that a rising RMSD indicates fibril dissociation and lack of stability, whereas a constant RMSD signifies a stable fibril structure. We assigned a RMSD ceiling of 3 Å, and when a fibril model's RMSD increased above this value it was judged to be dissociated. Figure D2 shows that the antiparallel and parallel β-sheet structures stayed well below the RMSD limit, while the β-hairpin model dissociated at ∼58 ns.

D.3 RESULTS AND DISCUSSION

D.3.1 UVRR of NDQ10 and DQ10 Fibril Films

Figure D3 shows the 197 nm – 204 nm UVRR difference spectra of dried NDQ10 and DQ10 fibril films. The AmI^P and $AmII^P$ bands are located at $\sim$1660 cm^{-1} and $\sim$1612 cm^{-1}, respectively, for both NDQ10 and DQ10. These bands negligibly shift compared to the AmI^P and $AmII^P$ bands of fibrils in solution, as shown in the Figure 5.3 197 nm – 204 nm UVRR difference spectra in the main text. This indicates that the hydrogen bonding environments of the Gln side chain primary amides are not significantly perturbed upon dehydrating the fibrils. This occurs because there is strong side chain inter-amide hydrogen bonding in NDQ10 and DQ10 fibrils.

The CH_2 wagging band is located at $\sim$1413 cm^{-1} for DQ10 fibril films and at $\sim$1410 cm^{-1} for NDQ10 fibril films. Compared to the Figure 5.3 197 nm – 204 nm UVRR difference spectra in the main text, the CH_2 wagging band downshifts $\sim$5 cm^{-1} for NDQ10 and $\sim$17 cm^{-1} for DQ10 fibrils upon dehydration. We attribute this $\sim$17 cm^{-1} downshift of the CH_2 wagging band in DQ10 fibrils upon dehdyration to a local dielectric environment change of the methylene groups around of the Gln side chains.

D.3.2 Band Assignments of Mono-deuterated Primary Amides

We employed DFT calculations to aid in our band assignments of the Gln UVRR spectrum in 50%/50% H_2O/D_2O (Figure 5.5 in the main text). In assigning the Figure 5.5 spectrum, we assume that vibrations containing significant contributions of $C_\delta - N_{\epsilon 2}$ stretching show resonance enhancement in the the UVRR spectrum because the electronic excited state is expected to be expanded along this coordinate [147]. Table D1 and Table D2 show the potential energy distributions (PEDs) obtained from the DFT calculations for the "*cis*-$N_{\epsilon 2}HD$" and "*trans*-$N_{\epsilon 2}HD$" species of the mono-deuterated primary amide side chains of Gln. Our band assignments of the Figure 5.5 spectrum from the main text are shown in Table D3.

D.3.2.1 Assignment of Amide Vibrations The DFT calculations show that partial deuteration of the Gln side chains results in a reorganization of the eigenvector composition of the primary amide vibrations compared to fully protonated side chains. This results in the decoupling of N–H and N–D motions and the appearance of vibrations that resemble canonical secondary amide modes. Our findings agree with Saito and coworkers' normal mode analyses [211, 212] of partially deuterated acetamide.

We assign the AmI vibration to a band located at $\sim$1660 cm^{-1}. Our normal mode analysis indicates that the PED of this vibration consists mostly of C_{δ}=$O_{\epsilon 1}$ stretching ($\sim$77%), but also contains significant contributions of C_{δ}–$N_{\epsilon 2}$ stretching and $N_{\epsilon 1}C_{\delta}C_{\gamma}$ bending. The PED of this vibration is essentially the same as that of the AmI^{S} and AmI^{P} vibrations [146, 180].

The DFT calculations also predict the appearance of $AmII^{S}$- and $AmIII^{S}$-like vibrations. Both peptide backbone C–N stretching and N–H in-plane bending motions are important in defining the PEDs of the canonical $AmII^{S}$ and $AmIII^{S}$ vibrations. Therefore, in the case of the mono-deuterated primary amides, we searched for vibrations that contain significant contributions of C_{δ}–$N_{\epsilon 2}$ stretching and $N_{\epsilon 2}$HD deformations. As shown in Table D1 and Table D2, there are several vibrations that contain significant contributions of C_{δ}–$N_{\epsilon 2}$ stretching, $N_{\epsilon 2}$HD scissoring, and $N_{\epsilon 2}$HD rocking. We assign these vibrations to $AmII^{S}$-like, $AmIII^{S}$-like, or $AmIII'^{S}$-like vibrations.

There are two $AmII^{S}$-like vibrations predicted by the DFT calculations to be at 1524 cm^{-1} for the *trans*-$N_{\epsilon 2}$HD species and $\sim$1479 cm^{-1} for the *cis*-$N_{\epsilon 2}$HD species. Both of these vibrations contain significant C_{δ}–$N_{\epsilon 2}$ stretching and $N_{\epsilon 2}$HD scissoring, although the 1524 cm^{-1} mode also contains $N_{\epsilon 2}$HD rocking. The predicted $\sim$1524 cm^{-1} mode is experimentally observed at $\sim$1547 cm^{-1}, while the predicted $\sim$1479 cm^{-1} vibration is observed at $\sim$1476 cm^{-1}.

The DFT calculations indicate that two $AmIII^{S}$-like vibrations for the *trans*-$N_{\epsilon 2}$HD species are predicted to occur at $\sim$1247 cm^{-1} and $\sim$1329 cm^{-1}. Both vibrations contain significant contributions of C_{δ}–$N_{\epsilon 2}$ stretching and $N_{\epsilon 2}$HD scissoring. However, as with the canonical $AmIII^{S}$ modes observed in peptides, these vibrations are significantly coupled since they contain significant contributions of CH_2 wagging and twisting, as well as C_{α}–H rocking [225]. We assign these $AmIII^{S}$-like vibrations to bands observed at $\sim$1247 cm^{-1} and $\sim$1308 cm^{-1}.

There are also two AmIII$'^S$-like vibrations predicted to be at $\sim$1055 cm^{-1} for the *trans*-N$_{\epsilon 2}$HD species and $\sim$953 cm^{-1} for the *cis*-N$_{\epsilon 2}$HD species. These vibrations resemble AmIII$'^S$-like modes since they both contain large contributions of N$_{\epsilon 2}$HD rocking. This is analogous to the canonical AmIII$'^S$, which is mostly N-D in-plane bending.

D.3.3 Bennett Acceptance Ratio Method

The Bennett acceptance ratio (BAR) is used here to estimate the free energy difference between two states. The full equation for the BAR is

$$\sum_{i=1}^{n_i} \frac{1}{1+\exp(\ln(n_i/n_j)+\beta\Delta U_{ij}-\beta\Delta G)} - \sum_{i=1}^{n_j} \frac{1}{1+\exp(\ln(n_j/n_i)+\beta\Delta U_{ji}-\beta\Delta G)} = 0 \tag{D.10}$$

where ΔG is the Gibbs free energy difference between states i and j (here antiparallel and parallel β-sheet fibrils), n_i, n_j are the number of samples used for states i and j, $\beta = k_B T$ (where k_B is the Boltzmann constant and T is the simulation temperature), and $\Delta U_{i,j}, U_{j,i}$ are the potential energy differences between states i and j. This equation is solved numerically using an iterative method.

Bennett clearly states that the best estimates of the free energy differences between states occurs when the extent of the energy overlap is greatest and when the density-of-states as a function of the energy difference is smoothest [226]. Here, we employ pymbar, which utilizes a multistate Bennett acceptance ratio method that can handle two or more states. Since we are working with two states, fibril models *a* and *b*, the multistate method is identical to the traditional BAR method derived for two states [207]. Figure D4 demonstrates the overlapping potential energy distributions necessary for a converged BAR calculation. To ensure comparable energetics, identical atom counts as well as system dimensions were used for all fibril simulations (see main text and NAMD configuration files for details).

D.3.4 Hydrogen Bonding Analysis

Figure D5 show the number of hydrogen bonds of fibril models *a* and *b* (from Figure 1.1 in the main text), which were obtained from the MD trajectories. Table D4 lists the average

number of the different categories of hydrogen bonds formed during the MD simulations for fibril models *a* and *b*. As disucussed in the main text, antiparallel β-sheet model *a* forms, on average, more peptide bond-peptide bond hydrogen bonds than the parallel β-sheet model *b*. In particular, the antiparallel β-sheet forms significantly more peptide bond-peptide bond hydrogen bonds than the parallel β-sheet. It is interesting to note that, if we assume a peptide bond-peptide bond hydrogen bond energy of $5\,\mathrm{kJ\,mol^{-1}}$, this hydrogen bonding difference would account for the majority of the $160.5\,\mathrm{kJ\,mol^{-1}}$ free energy difference between the antiparallel and parallel β-sheet fibril models.

Table D1: DFT Calculated Frequencies (cm^{-1}) and Assignments of *cis*-Glutamine-$N_{\epsilon 2}$HD. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

Assign.	Calc.	PED[a,b] (≥5% contribution)
ν_{as}COO	1715	$-\nu$C'O (53), νC'O (33), ρC_αC' (7)
$\delta_{as}NH_3$	1666	$-\delta_{as}$'NH_3 (48), $\delta_{as}NH_3$ (44), ρNH_3 (5)
$\delta_{as}NH_3$	1621	$-\delta_{as}NH_3$ (40), $-\delta_{as}$'NH_3 (40), $-\delta_s NH_3$ (10)
σCH_2	1494	$\sigma C_\beta H_2$ (87)
AmII-like	1479	$-\sigma$**$N_{\epsilon 2}$HD** (37), $-\nu$**$C_\delta N_{\epsilon 2}$** (28), $\nu C_\gamma C_\delta$ (10), $-\omega C_\gamma H_2$ (9), $\rho C_\delta O_{\epsilon 1}$ (6)
$\delta_s NH_3$	1464	$\delta_s NH_3$ (50), $-\sigma C_\gamma H_2$ (20)
σCH_2	1460	$\sigma C_\gamma H_2$ (61), $\delta_s NH_3$ (21)
ωCH_2	1437	$\omega C_\beta H_2$ (19), $\sigma N_{\epsilon 2}$HD (14), $-\omega C_\gamma H_2$ (11), $-\nu$C'O (10), ρC_αH (8), $-\nu C_\beta C_\gamma$ (5), $-\delta_s NH_3$ (5), $\sigma C_\gamma H_2$ (5)
ν_sCOO	1414	νC'O (22), $-\omega C_\gamma H_2$ (13), $\sigma N_{\epsilon 2}$HD (11), $-\nu C_\alpha$C' (9), $-\rho C_\alpha$H (7), $\nu C_\gamma C_\delta$ (7), $-\beta$C'O_2^- (6), νC'O (6)
ρCH	1387	ρC_αH (28), $\omega C_\beta H_2$ (11), $-\rho$'C_αH (9), νC'O (8), $\nu C_\alpha C_\beta$ (7), $-\tau C_\beta H_2$ (6)
ρCH	1357	ρC_αH (27), $\tau C_\beta H_2$ (19), $-\omega C_\beta H_2$ (13), $-\tau C_\gamma H_2$ (11)
$\omega CH_2+\tau CH_2$	1347	$\omega C_\beta H_2$ (21), $\tau C_\beta H_2$ (21), ρ'C_αH (11), ν**$C_\delta N_{\epsilon 2}$** (11), $-\rho C_\delta O_{\epsilon 1}$ (5)
τCH_2	1309	$\tau C_\gamma H_2$ (35), ρ'C_αH (26), $-\rho C_\beta H_2$ (8), ρC_αH (6)
ωCH_2	1272	$\omega C_\gamma H_2$ (43), $\omega C_\beta H_2$ (20), $-\nu$**$C_\delta N_{\epsilon 2}$** (13)
$\tau CH_2+\nu$CC	1214	$\tau C_\beta H_2$ (21), $\nu C_\alpha C_\beta$ (18), $\tau C_\gamma H_2$ (16), ρ'NH_3 (13), $-\delta$NC_αC'O_2^- (5)
$\rho CH+\tau CH_2$	1153	$-\rho$'C_αH (20), $\tau C_\gamma H_2$ (16), $-\rho$'NH_3 (14), $\tau C_\beta H_2$ (12), $-\nu C_\alpha C_\beta$ (8), $\rho C_\gamma H_2$ (6)
νCC	1117	$\nu C_\beta C_\gamma$ (42), νC_αN (10), ρNH_3 (8), $-\nu C_\alpha C_\beta$ (7), $-\delta$'NC_αC'O_2^- (5), $-\sigma C_\alpha C_\beta C_\gamma$ (5)
ρNH_3	1105	ρNH_3 (28), $-\nu C_\beta C_\gamma$ (12), $-\rho$'C_αH (9), $-\delta$'NC_αC'O_2^- (8), $-\rho C_\alpha$H (8), $-\rho C_\beta H_2$ (7)
νCN	1040	νC_αN (35), $-\nu C_\beta C_\gamma$ (11), $\rho C_\beta H_2$ (9), $\rho C_\gamma H_2$ (6), ρ'C_αH (5)
$\rho NH_3+\nu$CN	1005	ρNH_3 (22), $-\nu C_\alpha$N (20), $\rho C_\gamma H_2$ (12), $\rho C_\beta H_2$ (12), $\nu C_\alpha C_\beta$ (10)
ρNH_3	983	$-\rho$'NH_3 (29), $\nu C_\alpha C_\beta$ (13), $-\sigma C_\alpha C_\beta C_\gamma$ (6), $\rho N_{\epsilon 2}$HD (6), $-\sigma C_\beta C_\gamma C_\delta$ (5), $\nu C_\gamma C_\delta$ (5)
AmIII'[S]-like	953	ρ**$N_{\epsilon 2}$HD** (15), ν**$C_\delta N_{\epsilon 2}$** (14), $-\nu C_\alpha C_\beta$ (13), ρ'NH_3 (11), $\nu C_\gamma C_\delta$ (9), νC_αC' (8), $-\sigma N_{\epsilon 2}$HD (7)

[a] C_δ–$N_{\epsilon 2}$ stretching and $N_{\epsilon 2}$HD scissoring components in PED are in bold.

[b] ν: stretch; δ_{as}: asymmetric deformation; δ_s: symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

Table D2: DFT Calculated Frequencies (cm^{-1}) and Assignments of *trans*-Glutamine-$N_{\epsilon 2}$HD. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

Assign.	Calc.	PED[a,b] (≥5% contribution)
AmI-like	1742	$-\nu C_\delta O_{\epsilon 1}$ (77), $-\beta NC_\delta C_\gamma$ (7), $\nu \mathbf{C_\delta N_{\epsilon 2}}$ **(7)**
$\nu_{as}COO$	1715	$-\nu C'O$ (53), $\nu C'O$ (33), $\rho C_\alpha C'$ (7)
$\delta_{as}NH_3$	1666	$\delta_{as}'NH_3$ (48), $-\delta_{as}NH_3$ (44), $-\rho NH_3$ (5)
$\delta_{as}NH_3$	1621	$-\delta_{as}NH_3$ (40), $-\delta_{as}'NH_3$ (40), $-\delta_s NH_3$ (10)
AmII-like	1524	$\sigma \mathbf{N_{\epsilon 2}HD}$ **(49)**, $\nu \mathbf{C_\delta N_{\epsilon 2}}$ **(30)**, $-\nu C_\gamma C_\delta$ (6), $-\rho \mathbf{N_{\epsilon 2}HD}$ **(6)**, $-\rho C_\delta O_{\epsilon 1}$ (5)
σCH_2	1494	$\sigma C_\beta H_2$ (88)
$\delta_s NH_3+\sigma CH_2$	1465	$-\delta_s NH_3$ (36), $\sigma C_\gamma H_2$ (33), $\omega C_\gamma H_2$ (5)
σCH_2	1460	$-\sigma C_\gamma H_2$ (47), $-\delta_s NH_3$ (34), $\nu C'O$ (3)
ωCH_2	1435	$-\omega C_\beta H_2$ (21), $\nu C'O$ (13), $-\rho C_\alpha H$ (11), $\omega C_\gamma H_2$ (10), $-\sigma C_\gamma H_2$ (7), $-\nu C_\alpha C'$ (6), $\nu C_\beta C_\gamma$ (5), $\delta_s NH_3$ (5)
$\nu_s COO$	1407	$\nu C'O$ (22), $-\omega C_\gamma H_2$ (20), $-\nu C_\alpha C'$ (9), $\nu C_\gamma C_\delta$ (7), $-\beta C'O_2^-$ (6), $\nu C'O$ (6), $\sigma N_{\epsilon 2}HD$ (5)
ρCH	1384	$\rho C_\alpha H$ (40), $-\rho' C_\alpha H$ (11), $-\tau C_\beta H_2$ (6), $\nu C_\alpha C_\beta$ (6), $\nu C'O$ (5)
τCH_2	1356	$\tau C_\beta H_2$ (35), $\rho C_\alpha H$ (16), $-\tau C_\gamma H_2$ (15), $\rho' C_\alpha H$ (6)
$AmIII^S$-like	1329	$\omega C_\beta H_2$ (39), $\rho' C_\alpha H$ (9), $-\sigma \mathbf{N_{\epsilon 2}HD}$ **(9)**, $\nu \mathbf{C_\delta N_{\epsilon 2}}$ **(8)**, $-\nu C_\gamma C_\delta$ (5)
τCH_2	1307	$-\tau C_\gamma H_2$ (34), $-\rho' C_\alpha H$ (22), $\rho C_\beta H_2$ (8), $-\rho C_\alpha H$ (8)
$AmIII^S$-like	1247	$-\omega C_\gamma H_2$ (25), $\nu \mathbf{C_\delta N_{\epsilon 2}}$ (20), $-\sigma \mathbf{N_{\epsilon 2}HD}$ (9), $\rho' NH_3$ (7), $-\rho C_\delta O_{\epsilon 1}$ (6), $\nu C_\alpha C_\beta$ (6), $-\omega C_\beta H_2$ (5), $\tau C_\beta H_2$ (5)
τCH_2	1207	$-\tau C_\beta H_2$ (18), $-\tau C_\gamma H_2$ (15), $-\nu C_\alpha C_\beta$ (13), $-\omega C_\gamma H_2$ (11), $-\rho' NH_3$ (10), $\nu C_\delta N_{\epsilon 2}$ (7)
$\rho CH+\tau CH_2$	1153	$\rho' C_\alpha H$ (19), $-\tau C_\gamma H_2$ (16), $\rho' NH_3$ (14), $-\tau C_\beta H_2$ (11), $\nu C_\alpha C_\beta$ (9), $-\rho C_\gamma H_2$ (6)
νCN	1116	$-\nu C_\beta C_\gamma$ (39), $-\nu C_\alpha N$ (10), $-\rho NH_3$ (10), $\nu C_\alpha C_\beta$ (6), $\delta' NC_\alpha C'O_2^-$ (6), $\rho' C_\alpha H$ (5), $\sigma C_\alpha C_\beta C_\gamma$ (5)
ρNH_3	1104	$-\rho NH_3$ (27), $\nu C_\beta C_\gamma$ (15), $\rho' C_\alpha H$ (8), $\rho C_\alpha H$ (7), $\rho C_\beta H_2$ (7), $\delta' NC_\alpha C'O_2^-$ (7)
$AmIII'^S$-like	1055	$-\rho \mathbf{N_{\epsilon 2}HD}$ (27), $-\nu C_\gamma C_\delta$ (14), $-\nu \mathbf{C_\delta N_{\epsilon 2}}$ (9), $-\nu C_\alpha N$ (9), $-\nu C_\delta O_{\epsilon 1}$ (8), $\nu C_\beta C_\gamma$ (5), $-\rho C_\delta O_{\epsilon 1}$ (5)
νCN	1032	$\nu C_\alpha N$ (31), $-\rho N_{\epsilon 2}HD$ (10), $\rho' NH_3$ (7), $-\nu C_\beta C_\gamma$ (7), $\rho C_\beta H_2$ (6)
ρNH_3	1003	ρNH_3 (25), $-\nu C_\alpha N$ (19), $\rho C_\gamma H_2$ (14), $\rho C_\beta H_2$ (14), $\nu C_\alpha C_\beta$ (7), $-\tau C_\beta H_2$ (5)
ρNH_3	973	$-\rho' NH_3$ (37), $\nu C_\alpha C_\beta$ (25), $-\nu C_\alpha C'$ (9), $-\sigma C_\alpha C_\beta C_\gamma$ (7), $\nu C_\alpha N$ (6)

[a] C_δ–$N_{\epsilon 2}$ stretching and $N_{\epsilon 2}$HD scissoring components in PED are in bold.

[b] ν: stretch; δ_{as}: asymmetric deformation; δ_s: symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

Table D3: UVRR Band Frequencies (cm^{-1}) and Assignments of *trans*- and *cis*-Glutamine-$N_{\epsilon 2}HD$. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

cis-$N_{\epsilon 2}HD$			*trans*-$N_{\epsilon 2}HD$		
Expt.	Calc.	Assign.[a]	Expt.	Calc.	Assign.[a]
			1658	1742	AmI-like
1628	1715	$\nu_{as}COO$	1628	1715	$\nu_{as}COO$
	1666	$\delta_{as}NH_3$		1666	$\delta_{as}NH_3$
	1621	$\delta_{as}NH_3$		1621	$\delta_{as}NH_3$
			1547	1524	AmII-like
	1494	σCH_2		1494	σCH_2
1476	1479	AmII-like			
	1464	$\delta_s NH_3$		1465	$\delta_s NH_3+\sigma CH_2$
1449	1460	σCH_2	1449	1460	σCH_2
1420	1437	ωCH_2	1420	1435	ωCH_2
1420	1414	$\nu_s COO$	1420	1407	$\nu_{as}COO$
1398	1387	ρCH	1398	1384	ρCH
1360	1357	ρCH			
			1360	1356	τCH_2
1335	1347	$\omega CH_2+\tau CH_2$			
			1308	1329	$AmIII^S$-like
	1309	τCH_2		1307	τCH_2
1278	1272	ωCH_2			
			1247	1247	$AmIII^S$-like
	1214	$\tau CH_2+\nu CC$		1207	τCH_2
	1153	$\rho CH+\tau CH_2$	1153	1153	$\rho CH+\tau CH_2$
	1117	νCC		1116	νCC
	1105	ρNH_3		1104	ρNH_3
			1040	1055	$AmIII'^S$-like
	1040	νCN			
				1032	νCN
	1005	$\rho NH_3+\nu CN$		1003	ρNH_3
	983				ρNH_3
				973	ρNH_3
964	953	$AmIII'^S$-like			

[a] ν: stretch; δ_{as}: asymmetric deformation; δ_s: symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

Table D4: Average Number of Hydrogen Bonds for Antiparallel and Parallel β-sheet Fibril Models. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

H-bonding type	Parallel β-strand fibril	Antiparallel β-strand fibril
Peptide-peptide	106.7	118.9
Peptide-solvent	281.2	256.5
Peptide backbone-side chain	36.0	16.8
Peptide backbone-peptide backbone	39.2	66.8
Side chain-side chain	31.4	35.3

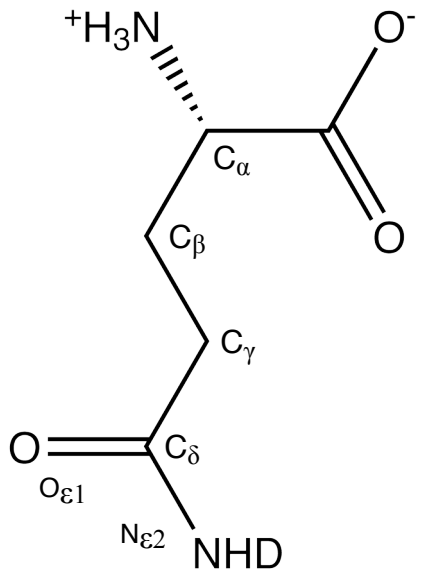


Figure D1: Atomic labeling scheme for glutamine used in DFT calculations. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

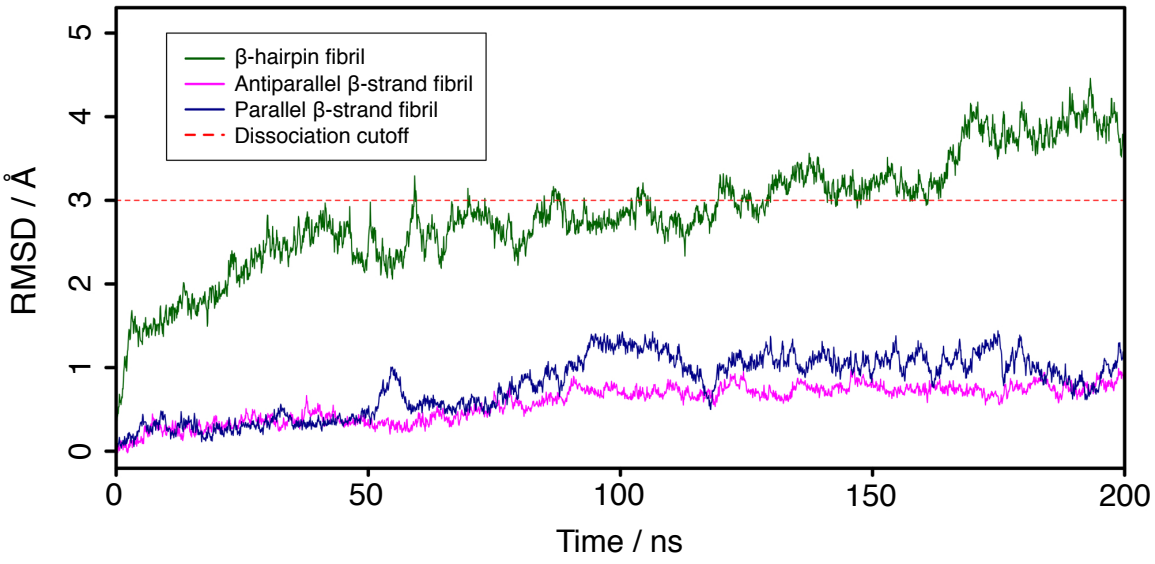


Figure D2: Plot of the evolution of backbone RMSD metric for three fibril models with respect to their initial structure. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

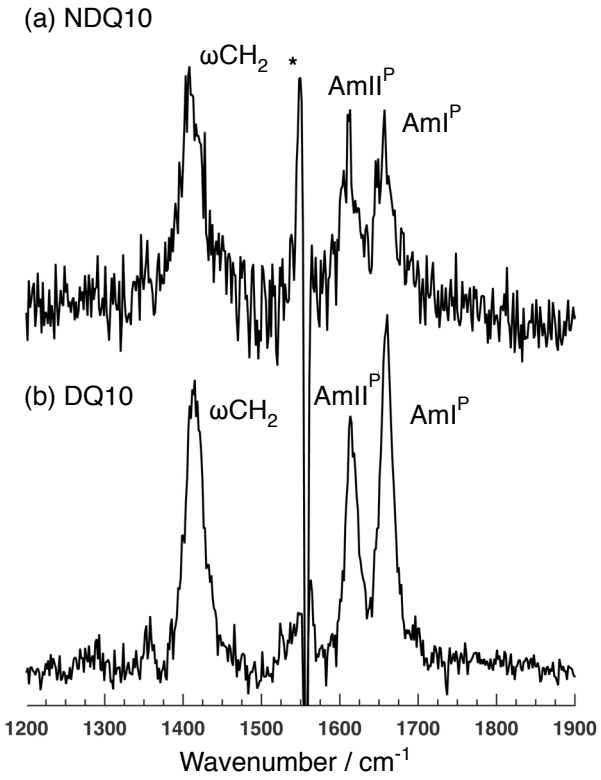


Figure D3: UVRR 197 nm – 204 nm difference spectra of fibril films prepared from (a) NDQ10 and (b) DQ10. The asterisk indicates an artifact of subtracting the intense O_2 stretching band in the difference spectrum. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

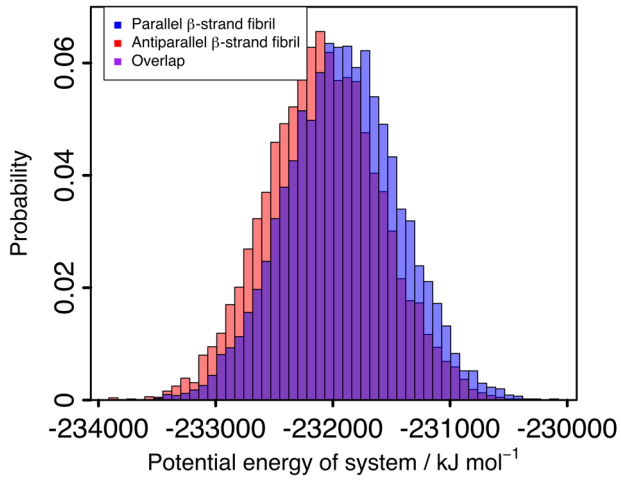


Figure D4: Potential energy distributions for the Figure 5.1 simulated model *a* antiparallel and model *b* parallel β-sheet fibril systems shown as red and blue histograms, respectively. These distributions represent the probability that a particular potential energy was sampled during the simulation for the antiparallel β-sheet fibril system (red) and the parallel β-sheet fibril system (blue). The significant overlap indicates that potential energies are sufficiently converged for use of the BAR method. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

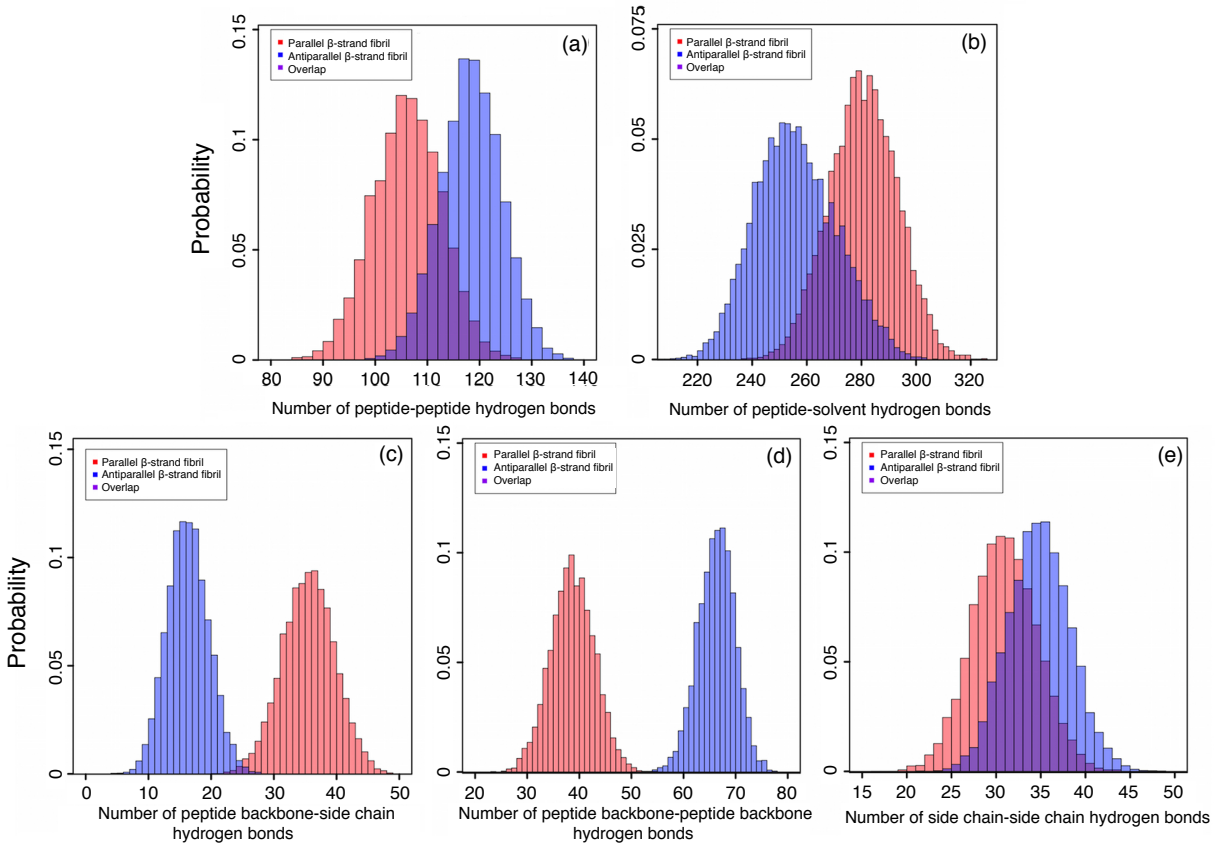


Figure D5: Histograms of the number of hydrogen bonds for models *a* and *b* in Figure 1.1. (a) peptide-peptide hydrogen bonds, (b) peptide-solvent hydrogen bonds, (c) peptide backbone-side chain hydrogen bonds, (d) peptide backbone-peptide backbone, and (e) side chain-side chain hydrogen bonds. Data was taken throughout the 200 ns trajectory for model *a* (blue) and model *b* (red). The overlap of the two distributions is shown in purple. Hydrogen bonds were defined as having a distance between heavy atoms of less than 3.0 Å and an angle cutoff of 30°. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

REFERENCES

[1] Harry T. Orr and Huda Y. Zoghbi. Trinucleotide Repeat Disorders. *Annual Review of Neuroscience*, 30(1):575–621, 2007.

[2] Russell L. Margolis, Dobrila D. Rudnicki, and Susan E. Holmesi. Huntington's Disease Like-2: Review and Update. *Acta Neurologica Taiwanica*, 14(1):1–8, 2005.

[3] Jennifer R. Gatchel and Huda Y. Zoghbi. Diseases of Unstable Repeat Expansion: Mechanisms and Common Principles. *Nature Reviews Genetics*, 6(10):743–755, 2007.

[4] Regina M. Murphy, Robert H. Walters, Matthew D. Tobelmann, and Joseph P. Bernacki. *Non-fibrillar Amyloidogenic Protein Assemblies - Common Cytotoxins Underlying Degenerative Diseases*, chapter When More Is Not Better: Expanded Polyglutamine Domains in Neurodegenerative Disease, pages 337–375. Springer Netherlands, Dordrecht, 2012.

[5] James F. Gusella and Marcy E. MacDonald. Molecular Genetics: Unmasking Polyglutamine Triggers in Neurodegenerative Disease. *Nature Reviews Neuroscience*, 1(2):109–115, 2000.

[6] Richard H. Myers. Huntington's Disease Genetics. *NeuroRX*, 1(2):255–262, 2004.

[7] Marian DiFiglia, Ellen Sapp, Kathryn O. Chase, Stephen W. Davies, Gillian P. Bates, J. P. Vonsattel, and Neil Aronin. Aggregation of Huntingtin in Neuronal Intranuclear Inclusions and Dystrophic Neurites in Brain. *Science*, 277(5334):1990–1993, 1997.

[8] A. Kazantsev, E. Preisinger, A. Dranovsky, D. Goldgaber, and D. Housman. Insoluble Detergent-Resistant Aggregates form Between Pathological and Nonpathological Lengths of Polyglutamine in Mammalian Cells. *Proceedings of the National Academy of Sciences*, 96(20):11404–11409, 1999.

[9] Joan S. Steffan, Aleksey Kazantsev, Olivera Spasic-Boskovic, Marilee Greenwald, Ya-Zhen Zhu, Heike Gohler, Erich E. Wanker, Gillian P. Bates, David E. Housman, and Leslie M. Thompson. The Huntington's Disease Protein Interacts with P53 and CREB-Binding Protein and Represses Transcription. *Proceedings of the National Academy of Sciences*, 97(12):6763–6768, 2000.

[10] Joan S. Steffan, Laszlo Bodai, Judit Pallos, Marnix Poelman, Alexander McCampbell, Barbara L. Apostol, Alexsey Kazantsev, Emily Schmidt, Ya-Zhen Zhu, Marilee Greenwald, et al. Histone Deacetylase Inhibitors Arrest Polyglutamine-Dependent Neurodegeneration in Drosophila. *Nature*, 413(6857):739–743, 2001.

[11] Sebastien Holbert, Isabelle Denghien, Tamara Kiechle, Adam Rosenblatt, Cheryll Wellington, Michael R. Hayden, Russell L. Margolis, Christopher A. Ross, Jean Dausset, Robert J. Ferrante, et al. The Gln-Ala Repeat Transcriptional Activator CA150 Interacts with Huntingtin: Neuropathologic and Genetic Evidence for a Role in Huntington's Disease Pathogenesis. *Proceedings of the National Academy of Sciences*, 98(4):1811–1816, 2001.

[12] Frederick C. Nucifora, Masayuki Sasaki, Matthew F. Peters, Hui Huang, Jillian K. Cooper, Mitsunori Yamada, Hitoshi Takahashi, Shoji Tsuji, Juan Troncoso, Valina L. Dawson, et al. Interference by Huntingtin and Atrophin-1 with CBP-Mediated Transcription Leading to Cellular Toxicity. *Science*, 291(5512):2423–2428, 2001.

[13] Nihar Ranjan Jana, Motomasa Tanaka, Guang-hui Wang, and Nobuyuki Nukina. Polyglutamine Length-Dependent Interaction of Hsp40 and Hsp70 Family Chaperones with Truncated N-terminal Huntingtin: their Role in Suppression of Aggregation and Cellular Toxicity. *Human Molecular Genetics*, 9(13):2009–2018, 2000.

[14] Yaohui Chai, Stacia L. Koppenhafer, Sarah J. Shoesmith, Matthew K. Perez, and Henry L. Paulson. Evidence for Proteasome Involvement in Polyglutamine Disease: Localization to Nuclear Inclusions in SCA3/MJD and Suppression of Polyglutamine Aggregation In Vitro. *Human Molecular Genetics*, 8(4):673–682, 1999.

[15] Andrej Michalik and Christine Van Broeckhoven. Pathogenesis of Polyglutamine Disorders: Aggregation Revisited. *Human Molecular Genetics*, 12(suppl 2):R173–R186, 2003.

[16] H. Diana Rosas, David H. Salat, Stephanie Y. Lee, Alexandra K. Zaleta, Nathanael Hevelone, and Steven M. Hersch. Complexity and Heterogeneity: What Drives the Ever-changing Brain in Huntington's disease? *Annals of the New York Academy of Sciences*, 1147(1):196–205, 2008.

[17] Wen Yang, John R. Dunlap, Richard B. Andrews, and Ronald Wetzel. Aggregated Polyglutamine Peptides Delivered to Nuclei are Toxic to Mammalian Cells. *Human Molecular Genetics*, 11(23):2905–2917, 2002.

[18] Toshiya Sato, Masami Miura, Mitsunori Yamada, Takayuki Yoshida, Jonathan D. Wood, Ikuru Yazawa, Masao Masuda, Takeo Suzuki, Ryong-Moon Shin, Hau-Jie Yau, et al. Severe Neurological Phenotypes of Q129 DRPLA Transgenic Mice Serendipitously Created by En Masse Expansion of CAG Repeats in Q76 DRPLA Mice. *Human Molecular Genetics*, 18(4):723–736, 2009.

[19] Claire-Anne Gutekunst, Shi-Hua Li, Hong Yi, James S. Mulroy, Stefan Kuemmerle, Randi Jones, David Rye, Robert J. Ferrante, Steven M. Hersch, and Xiao-Jiang Li. Nuclear and Neuropil Aggregates in Huntingtons Disease: Relationship to Neuropathology. *The Journal of Neuroscience*, 19(7):2522–2534, 1999.

[20] Estrella Gomez-Tortosa, Marcy E. MacDonald, Julia C. Friend, Taylor Sherryl A.M., Larry J. Weiler, L. Adrienne Cupples, Jayalakshmi Srinidhi, James F. Gusella, Edward D. Bird, Jean-Paul Vonsattel, et al. Quantitative Neuropathological Changes in Presymptomatic Huntington's Disease. *Annals of Neurology*, 49(1):29–34, 2001.

[21] Christopher J. Cummings, Michael A. Mancini, Barbara Antalffy, Donald B. DeFranco, Harry T. Orr, and Huda Y. Zoghbi. Chaperone Suppression of Aggregation and Altered Subcellular Proteasome Localization Imply Protein Misfolding in SCA1. *Nature Genetics*, 19(2):148–154, 1998.

[22] Ashwani K. Thakur, Wen Yang, and Ronald Wetzel. Inhibition of Polyglutamine Aggregate Cytotoxicity by a Structure-Based Elongation Inhibitor. *The FASEB Journal*, 18(7):923–925, 2004.

[23] Frdric Saudou, Steven Finkbeiner, Didier Devys, and Michael E. Greenberg. Huntingtin Acts in the Nucleus to Induce Apoptosis but Death Does Not Correlate with the Formation of Intranuclear Inclusions. *Cell*, 95(1):55 – 66, 1998.

[24] Pratibha Siwach, Sonali Sengupta, Rashmi Parihar, and Subramaniam Ganesh. Spatial Positions of Homopolymeric Repeats in the Human Proteome and their Effect on Cellular Toxicity. *Biochemical and Biophysical Research Communications*, 380(2):382 – 386, 2009.

[25] Elizabeth J. Slow, Rona K. Graham, Alexander P. Osmand, Rebecca S. Devon, Ge Lu, Yu Deng, Jacqui Pearson, Kuljeet Vaid, Nagat Bissada, Ronald Wetzel, Blair R. Leavitt, et al. Absence of Behavioral Abnormalities and Neurodegeneration In Vivo Despite Widespread Neuronal Huntingtin Inclusions. 102(32):11402–11407, 2005.

[26] Christopher J. Cummings, Eyal Reinstein, Yaling Sun, Barbara Antalffy, Yong-hui Jiang, Aaron Ciechanover, Harry T. Orr, Arthur L. Beaudet, and Huda Y. Zoghbi. Mutation of the E6-AP Ubiquitin Ligase Reduces Nuclear Inclusion Frequency while Accelerating Polyglutamine-Induced Pathology in SCA1 Mice. *Neuron*, 24(4):879–892, 1999.

[27] Seung-Yun Yoo, Mark E. Pennesi, Edwin J. Weeber, Bisong Xu, Richard Atkinson, Shiming Chen, Dawna L. Armstrong, Samuel M. Wu, J. David Sweatt, and Huda Y. Zoghbi. SCA7 Knockin Mice Model Human SCA7 and Reveal Gradual Accumulation of Mutant Ataxin-7 in Neurons and Abnormalities in Short-Term Plasticity. *Neuron*, 37(3):383 – 401, 2003.

[28] Kei Watase, Edwin J. Weeber, Bisong Xu, Barbara Antalffy, Lisa Yuva-Paylor, Kouichi Hashimoto, Masanobu Kano, Richard Atkinson, Yaling Sun, Dawna L. Armstrong,

et al. A Long CAG Repeat in the Mouse Sca1 Locus Replicates SCA1 Features and Reveals the Impact of Protein Solubility on Selective Neurodegeneration. *Neuron*, 34(6):905 – 919, 2002.

[29] Montserrat Arrasate, Siddhartha Mitra, Erik S. Schweitzer, Mark R. Segal, and Steven Finkbeiner. Inclusion Body Formation Reduces Levels of Mutant Huntingtin and the Risk of Neuronal Death. *Nature*, 431(7010):805–810, 2004.

[30] Yoko Nekooki-Machida, Masaru Kurosawa, Nobuyuki Nukina, Kazuki Ito, Toshiro Oda, and Motomasa Tanaka. Distinct conformations of in vitro and in vivo amyloids of huntingtin-exon1 show different cytotoxicity. *Proceedings of the National Academy of Sciences*, 106(24):9679–9684, 2009.

[31] E.L. Altschuler, N.V. Hud, J.A. Mazrimas, and B. Rupp. Random Coil Conformation for Extended Polyglutamine Stretches in Aqueous Soluble Monomeric Peptides. *Journal of Peptide Research*, 50(1):73–75, 1997.

[32] Songming Chen, Valerie Berthelier, J. Bradley Hamilton, Brian O'Nuallain, and Ronald Wetzel. Amyloid-like Features of Polyglutamine Aggregates and Their Assembly Kinetics. *Biochemistry*, 41(23):7391–7399, 2002.

[33] Brian W. Chellgren, Anne-Frances Miller, and Trevor P. Creamer. Evidence for Polyproline II Helical Structure in Short Polyglutamine Tracts. *Journal of Molecular Biology*, 361(2):362 – 371, 2006.

[34] Fabrice A.C. Klein, Annalisa Pastore, Laura Masino, Gabrielle Zeder-Lutz, Helene Nierengarten, Mustapha Oulad-Abdelghani, Danile Altschuh, Jean-Louis Mandel, and Yvon Trottier. Pathogenic and Non-pathogenic Polyglutamine Tracts Have Similar Structural Properties: Towards a Length-dependent Toxicity Gradient. *Journal of Molecular Biology*, 371(1):235 – 244, 2007.

[35] Ronald Wetzel. Physical Chemistry of Polyglutamine: Intriguing Tales of a Monotonous Sequence. *Journal of Molecular Biology*, 421(4-5):466–90, 2012.

[36] Xiaoling Wang, Andreas Vitalis, Matthew A. Wyczalkowski, and Rohit V. Pappu. Characterizing the Conformational Ensemble of Monomeric Polyglutamine. *Proteins*, 63(2):297–311, 2006.

[37] Andreas Vitalis, Xiaoling Wang, and Rohit V. Pappu. Atomistic Simulations of the Effects of Polyglutamine Chain Length and Solvent Quality on Conformational Equilibria and Spontaneous Homodimerization. *Journal of Molecular Biology*, 384(1):279 – 297, 2008.

[38] Andreas Vitalis, Nicholas Lyle, and Rohit V. Pappu. Thermodynamics of β-Sheet Formation in Polyglutamine. *Biophysical Journal*, 97(1):303 – 311, 2009.

[39] Miki Nakano, Hirofumi Watanabe, Stuart M. Rothstein, and Shigenori Tanaka. Comparative Characterization of Short Monomeric Polyglutamine Peptides by Replica Exchange Molecular Dynamics Simulation. *The Journal of Physical Chemistry B*, 114(20):7056–7061, 2010.

[40] Scott L. Crick, Murali Jayaraman, Carl Frieden, Ronald Wetzel, and Rohit V. Pappu. Fluorescence Correlation Spectroscopy Shows that Monomeric Polyglutamine Molecules form Collapsed Structures in Aqueous Solutions. *Proceedings of the National Academy of Sciences*, 103(45):16764–16769, 2006.

[41] Andreas Vitalis, Xiaoling Wang, and Rohit V. Pappu. Quantitative Characterization of Intrinsic Disorder in Polyglutamine: Insights from Analysis Based on Polymer Theories. *Biophysical Journal*, 93(6):1923 – 1937, 2007.

[42] Vijay R. Singh and Lisa J. Lapidus. The Intrinsic Stiffness of Polyglutamine Peptides. *The Journal of Physical Chemistry B*, 112(42):13172–13176, 2008.

[43] Robert H. Walters and Regina M. Murphy. Examining Polyglutamine Peptide Length: A Connection between Collapsed Conformations and Increased Aggregation. *Journal of Molecular Biology*, 393(4):978 – 992, 2009.

[44] Andreas Vitalis and Rohit V. Pappu. Assessing the Contribution of Heterogeneous Distributions of Oligomers to Aggregation Mechanisms of Polyglutamine Peptides. *Biophys. Chem.*, 159(1):14 – 23, 2011.

[45] Yvon Trottier, Yves Lutz, Giovann Stevanin, Georges Imbert, Didier Devys, Geraldine Cancel, Frederic Saudou, Chantal Weber, Gilles David, Laszlo Tora, Yves Agid, Alexis Brice, and Jean-Louis Mandel. Polyglutamine Expansion as a Pathological Epitope in Huntington's disease and Four Dominant Cerebellar Ataxias. *Nature*, 378(6555):403–406, 1995.

[46] Jason Miller, Montserrat Arrasate, Elizabeth Brooks, Clare Peters Libeu, Justin Legleiter, Danny Hatters, Jessica Curtis, Kenneth Cheung, Preethi Krishnan, Siddhartha Mitra, et al. Identifying Polyglutamine Protein Species In Situ that Best Predict Neurodegeneration. *Nat. Chem. Bio.*, 7(12):925–934, 2011.

[47] Clare Peters-Libeu, Jason Miller, Earl Rutenber, Yvonne Newhouse, Preethi Krishnan, Kenneth Cheung, Danny Hatters, Elizabeth Brooks, Kartika Widjaja, Tina Tran, et al. Disease-Associated Polyglutamine Stretches in Monomeric Huntingtin Adopt a Compact Structure. *Journal of Molecular Biology*, 421(45):587 – 600, 2012. Amyloid Structure, Function, and Molecular Mechanisms (Part II).

[48] Melanie J. Bennett, Kathryn E. Huey-Tubman, Andrew B. Herr, Anthony P. West, Scott A. Ross, and Pamela J. Bjorkman. A Linear Lattice Model for Polyglutamine in CAG-Expansion Diseases. *Proceedings of the National Academy of Sciences*, 99(18):11634–11639, 2002.

[49] Pingwei Li, Kathryn E. Huey-Tubman, Tiyu Gao, Xiaojun Li, Anthony P. West, Melanie J. Bennett, and Pamela J. Bjorkman. The Structure of a PolyQ–Anti-PolyQ Complex RevealsBbinding According to a Linear Lattice Model. *Nature Structural and Molecular Biology*, 14(5):381–387, 2007.

[50] Louise C. Serpell. Alzheimer's Amyloid Fibrils: Structure and Assembly. *Biochimica et Biophysica Acta–Molecular Basis of Disease*, 1502(1):16 – 30, 2000.

[51] Markus S. Miettinen, Volker Knecht, Luca Monticelli, and Zoya Ignatova. Assessing Polyglutamine Conformation in the Nucleating Event by Molecular Dynamics Simulations. *The Journal of Physical Chemistry B*, 116(34):10259–10265, 2012.

[52] M F Perutz, T Johnson, M Suzuk, and J T Finch. Glutamine Repeats as Polar Zippers: Their Possible Role in Inherited Neurodegenerative Diseases. *Proceedings of the National Academy of Sciences*, 91(12):5355–5358, 1994.

[53] M. F. Perutz, J. T. Finch, J. Berriman, and A. Lesk. Amyloid Fibers are Water-Filled Nanotubes. *Proceedings of the National Academy of Sciences*, 99(8):5591–5595, 2002.

[54] Pawel Sikorski and Edward Atkins. New Model for Crystalline Polyglutamine Assemblies and their Connection with Amyloid Fibrils. *Biomacromolecules*, 6(1):425–432, 2005.

[55] Deepak Sharma, Leonid M. Shinchuk, Hideyo Inouye, Ronald Wetzel, and Daniel A. Kirschner. Polyglutamine Homopolymers Having 8–45 Residues Form Slablike β-crystallite Assemblies. *Proteins*, 61(2):398–411, 2005.

[56] Robert Schneider, Miria C. Schumacher, Henrik Mueller, Deepak Nand, Volker Klaukien, Henrike Heise, Dietmar Riedel, Gerhard Wolf, Elmar Behrmann, Stefan Raunser, and et al. Structural Characterization of Polyglutamine Fibrils by Solid-State NMR Spectroscopy. *Journal of Molecular Biology*, 412(1):121–136, 2011.

[57] Aneta T. Petkova, Yoshitaka Ishii, John J. Balbach, Oleg N. Antzutkin, Richard D. Leapman, Frank Delaglio, and Robert Tycko. A Structural Model for Alzheimer's β-amyloid Fibrils Based on Experimental Constraints from Solid State NMR. *Proceedings of the National Academy of Sciences*, 99(26):16742–16747, 2002.

[58] V. N. Sivanandam, Murali Jayaraman, Cody L. Hoop, Ravindra Kodali, Ronald Wetzel, and Patrick C. A. van der Wel. The Aggregation-Enhancing Huntingtin N-Terminus Is Helical in Amyloid Fibrils. *Journal of the American Chemical Society*, 133(12):4558–4566, 2011.

[59] Karunakar Kar, Cody L. Hoop, Kenneth W. Drombosky, Matthew A. Baker, Ravindra Kodali, Irene Arduini, Patrick C.A. van der Wel, W. Seth Horne, and Ronald Wetzel. β-Hairpin-Mediated Nucleation of Polyglutamine Amyloid Formation. *Journal of Molecular Biology*, 425(7):1183 – 1197, 2013.

[60] Cody L. Hoop, Hsiang-Kai Lin, Karunakar Kar, Zhipeng Hou, Michelle A. Poirier, Ronald Wetzel, and Patrick C. A. van der Wel. Polyglutamine Amyloid Core Boundaries and Flanking Domain Dynamics in Huntingtin Fragment Fibrils Determined by Solid-State Nuclear Magnetic Resonance. *Biochemistry*, 53(42):6653–6666, 2014.

[61] C. L. Hoop, H-K. Lin, K. Kar, G. Magyarfalvi, J.M. Lamley, J.C. Boatz, A. Mandal, Wetzel R. Lewandowski, J.R., and P.C.A. van der Wel. Huntingtin Exon 1 Fibrils Feature an Interdigitated β-hairpin-based Polyglutamine Core. *Proceedings of the National Academy of Sciences*, 113(6):1546–1551, 2016.

[62] Alexander J. Marchut and Carol K. Hall. Effects of Chain Length on the Aggregation of Model Polyglutamine Peptides: Molecular Dynamics Simulations. *Proteins*, 66(1):96–109, 2007.

[63] Luciana Esposito, Antonella Paladino, Carlo Pedone, and Luigi Vitagliano. Insights into Structure, Stability, and Toxicity of Monomeric and Aggregated Polyglutamine Models from Molecular Dynamics Simulations. *Biophysical Journal*, 94(10):4031 – 4040, 2008.

[64] Giulia Rossetti, Alessandra Magistrato, Annalisa Pastore, Francesca Persichetti, and Paolo Carloni. Structural Properties of Polyglutamine Aggregates Investigated via Molecular Dynamics Simulations. *The Journal of Physical Chemistry B*, 112(51):16843–16850, 2008.

[65] Rozita Laghaei and Normand Mousseau. Spontaneous Formation of Polyglutamine Nanotubes with Molecular Dynamics Simulations. *The Journal of Chemical Physics*, 132(16), 2010.

[66] Markus S. Miettinen, Luca Monticelli, Praveen Nedumpully-Govindan, Volker Knecht, and Zoya Ignatova. Stable Polyglutamine Dimers Can Contain β-Hairpins with Interdigitated Side Chains–But Not α-Helices, β-Nanotubes, β-Pseudohelices, or Steric Zippers. *Biophysical Journal*, 106(8):1721 – 1728, 2014.

[67] Songming Chen, Valerie Berthelier, Wen Yang, and Ronald Wetzel. Polyglutamine Aggregation Behavior In Vitro Supports a Recruitment Mechanism of Cytotoxicity. *Journal of Molecular Biology*, 311(1):173 – 182, 2001.

[68] Songming Chen, Frank A. Ferrone, and Ronald Wetzel. Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proceedings of the National Academy of Sciences*, 99(18):11884–11889, 2002.

[69] Ashwani K. Thakur and Ronald Wetzel. Mutational Analysis of the Structural Organization of Polyglutamine Aggregates. *Proceedings of the National Academy of Sciences*, 99(26):17014–17019, 2002.

[70] Karunakar Kar, Murali Jayaraman, Bankanidhi Sahoo, Ravindra Kodali, and Ronald Wetzel. Critical Nucleus Size for Disease-Related Polyglutamine Aggregation is Repeat-Length Dependent. *Nature Structural and Molecular Biology*, 18(3):328–336, 2011.

[71] Elizabeth Landrum and Ronald Wetzel. Biophysical Underpinnings of the Repeat Length Dependence of Polyglutamine Amyloid Formation. *The Journal of Biological Chemistry*, 289(15):10254–10260, 2014.

[72] Brian O'Nuallain, Ashwani K. Thakur, Angela D. Williams, Anusri M. Bhattacharyya, Songming Chen, Geetha Thiagarajan, and Ronald Wetzel. Kinetics and Thermodynamics of Amyloid Assembly Using a High Performance Liquid ChromatographyBased Sedimentation Assay. In *Amyloid, Prions, and Other Protein Aggregates, Part C*, volume 413 of *Methods in Enzymology*, pages 34 – 74. Academic Press, 2006.

[73] Anusri M. Bhattacharyya, Ashwani K. Thakur, and Ronald Wetzel. Polyglutamine aggregation nucleation: Thermodynamics of a highly unfavorable protein folding reaction. *Proceedings of the National Academy of Sciences*, 102(43):15400–15405, 2005.

[74] Joseph P. Bernacki and Regina M. Murphy. Model Discrimination and Mechanistic Interpretation of Kinetic Data in Protein Aggregation Studies. *Biophysical Journal*, 96(7):2871 – 2887, 2009.

[75] Christine C. Lee, Robert H. Walters, and Regina M. Murphy. Reconsidering the Mechanism of Polyglutamine Peptide Aggregation. *Biochemistry*, 46(44):12810–12820, 2007.

[76] Michael G. Zagorski, Jing Yang, Haiyan Shao, Kan Ma, Hong Zeng, and Anita Hong. Methodological and Chemical Factors Affecting Amyloid β Peptide Amyloidogenicity. In *Amyloid, Prions, and Other Protein Aggregates*, volume 309 of *Methods in Enzymology*, pages 189 – 204. Academic Press, 1999.

[77] Songming Chen and Ronald Wetzel. Solubilization and Disaggregation of Polyglutamine Peptides. *Protein Science*, 10(4):887–891, 2001.

[78] Shae B. Padrick, , and Andrew D. Miranker. Islet Amyloid: Phase Partitioning and Secondary Nucleation Are Central to the Mechanism of Fibrillogenesis. *Biochemistry*, 41(14):4694–4703, 2002.

[79] Michael R. Nichols, Melissa A. Moss, Dana Kim Reed, Stephanie Cratic-McDaniel, Jan H. Hoh, and Terrone L. Rosenberry. Amyloid-β Protofibrils Differ from Amyloid-β Aggregates Induced in Dilute Hexafluoroisopropanol in Stability and Morphology. *Journal of Biological Chemistry*, 280(4):2471–2480, 2005.

[80] Christopher P. Jaroniec, Cait E. MacPhee, Vikram S. Bajaj, Michael T. McMahon, Christopher M. Dobson, and Robert G. Griffin. High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. *Proceedings of the National Academy of Sciences*, 101(3):711–716, 2004.

[81] Anant K. Paravastu, Richard D. Leapman, Wai-Ming Yau, and Robert Tycko. Molecular Structural Basis for Polymorphism in Alzheimer's β-Amyloid Fibrils. *Proceedings of the National Academy of Sciences*, 105(47):18349–18354, 2008.

[82] Christian Wasmer, Adam Lange, Hélène Van Melckebeke, Ansgar B Siemer, Roland Riek, and Beat H Meier. Amyloid Fibrils of the HET-s (218-289) Prion Form a β Solenoid with a Triangular Hydrophobic Core. *Science*, 319(5869):1523–1526, 2008.

[83] Robert Tycko, Kimberly L. Sciarretta, Joseph P.R.O. Orgel, and Stephen C. Meredith. Evidence for Novel β-Sheet Structures in Iowa Mutant β-Amyloid Fibrils. *Biochemistry*, 48(26):6072–6084, 2009.

[84] Robert Tycko. BIOMOLECULAR SOLID STATE NMR: Advances in Structural Methodology and Applications to Peptide and Protein Fibrils. *Annual Review of Physical Chemistry*, 52(1):575–606, 2001.

[85] Robert Tycko. Progress Towards a Molecular-Level Structural Understanding of Amyloid Fibrils. *Current Opinion in Structural Biology*, 14(1):96 – 103, 2004.

[86] Robert Tycko. Molecular Structure of Amyloid Fibrils: Insights from Solid-State NMR. *Quarterly Reviews of Biophysics*, 39:1–55, 2 2006.

[87] Robert Tycko. Solid-State NMR Studies of Amyloid Fibril Structure. *Annual Review of Physical Chemistry*, 62(1):279–299, 2011.

[88] Galia T. Debelouchina, Marvin J. Bayro, Patrick C. A. van der Wel, Marc A. Caporini, Alexander B. Barnes, Melanie Rosay, Werner E. Maas, and Robert G. Griffin. Dynamic Nuclear Polarization-Enhanced Solid-State NMR Spectroscopy of GNNQQNY Nanocrystals and Amyloid Fibrils. *Physical Chemistry Chemical Physics*, 12:5911–5919, 2010.

[89] Sulayman A. Oladepo, Kan Xiong, Zhenmin Hong, Sanford A. Asher, Joseph Handen, and Igor K. Lednev. UV Resonance Raman Investigations of Peptide and Protein Structure and Dynamics. *Chemical Reviews*, 112(5):2604–2628, 2012.

[90] Ming Xu, Victor A. Shashilov, Vladimir V. Ermolenkov, Laura Fredriksen, Dmitri Zagorevski, and Igor K. Lednev. The First step of Hen Egg White Lysozyme Fibrillation, Irreversible Partial Unfolding, is a Two-State Transition. *Protein Science*, 16(5):815–832, 2007.

[91] Ming Xu, Victor Shashilov, and Igor K. Lednev. Probing the Cross-β Core Structure of Amyloid Fibrils by Hydrogen-Deuterium Exchange Deep Ultraviolet Resonance Raman Spectroscopy. *Journal of the American Chemical Society*, 129(36):11002–11003, 2007.

[92] Victor Shashilov, Ming Xu, Vladimir V. Ermolenkov, Laura Fredriksen, , and Igor K. Lednev. Probing a Fibrillation Nucleus Directly by Deep Ultraviolet Raman Spectroscopy. *Journal of the American Chemical Society*, 129(22):6972–6973, 2007.

[93] Victor A. Shashilov andIgor K. Lednev. 2D Correlation Deep UV Resonance Raman Spectroscopy of Early Events of Lysozyme Fibrillation: Kinetic Mechanism and Potential Interpretation Pitfalls. *Journal of the American Chemical Society*, 130(1):309–317, 2008.

[94] Ludmila A. Popova, Ravindra Kodali, Ronald Wetzel, and Igor K. Lednev. Structural Variations in the Cross-β Core of Amyloid β Fibrils Revealed by Deep UV Resonance Raman Spectroscopy. *Journal of the American Chemical Society*, 132(18):6324–6328, 2010.

[95] Dmitry Kurouski, William Lauro, and Igor K. Lednev. Amyloid Fibrils are "Alive": Spontaneous Refolding from One Polymorph to Another. *Chemical Communications*, 46:4249–4251, 2010.

[96] Sandorf A. Asher, Anatoli Ianoul, Guido Mix, Mary N. Boyden, Anton Karnoup, Max Diem, and Reinhard Schweitzer-Stenner. Dihedral ψ Angle Dependence of the Amide III Vibration: a Uniquely Sensitive UV Resonance Raman Secondary Structural Probe. *Journal of the American Chemical Society*, 123(47):11775–11781, 2001.

[97] Aleksandr V. Mikhonin, Sergei V. Bykov, Nataliya S. Myshakina, and Sanford A. Asher. Peptide Secondary Structure Folding Reaction Coordinate: Correlation Between UV Raman Amide III Frequency, ψ Ramachandran Angle, and Hydrogen Bonding. *The Journal of Physical Chemistry B*, 110(4):1928–1943, 2006.

[98] Nataliya S. Myshakina, Zeeshan Ahmed, and Sanford A. Asher. Dependence of Amide Vibrations on Hydrogen Bonding. *The Journal of Physical Chemistry B*, 112(38):11873–11877, 2008.

[99] Sanford A. Asher, Peter J. Larkin, and Junji Teraoka. Ultraviolet Resonance Raman and Absorption difference Spectroscopy of Myoglobins: Titration Behavior of Individual Tyrosine Residues. *Biochemistry*, 30(24):5944–5954, 1991.

[100] Peter J. Larkin, William G. Gustafson, and Sanford A. Asher. A New Raman Cross Section Measurement Technique Monitors the Tyrosine Environmental Dependence of the Electromagnetic Field Strength. *The Journal of Physical Chemistry*, 94(8):5324–5330, 1991.

[101] Zhenhuan Chi and Sanford A. Asher. UV Raman Determination of the Environment and Solvent Exposure of Tyr and Trp Residues. *The Journal of Physical Chemistry B*, 102(47):9595–9602, 1998.

[102] Zeeshan Ahmed, Nataliya S. Myshakina, and Sanford A. Asher. Dependence of the AmII'p Proline Raman Band on Peptide Conformation. *The Journal of Physical Chemistry B*, 113(32):11252–11259, 2009.

[103] Zhenmin Hong, Jonathan Wert, and Sanford A. Asher. UV Resonance Raman and DFT Studies of Arginine Side Chains in Peptides: Insights into Arginine Hydration. *The Journal of Physical Chemistry B*, 117(24):7145–7156, 2013.

[104] Igor K. Lednev, Anton S. Karnoup, Mark C. Sparrow, and Sanford A. Asher. α-Helix Peptide Folding and Unfolding Activation Barriers: A Nanosecond UV Resonance Raman Study. *Journal of the American Chemical Society*, 121(35):8074–8086, 1999.

[105] Sanford A. Asher, Alexander V. Mikhonin, and Sergei Bykov. UV Raman Demonstrates that α-Helical Polyalanine Peptides Melt to Polyproline II Conformations. *Journal of the American Chemical Society*, 126(27):8433–8440, 2004.

[106] Aleksandr V. Mikhonin and Sanford A. Asher. Direct UV Raman Monitoring of 3_{10}-Helix and π-Bulge Premelting During α-Helix Unfolding. *Journal of the American Chemical Society*, 128(42):13789–13795, 2006.

[107] Lu Ma, Zhenmin Hong, Bhavya Sharma, and Sanford A Asher. UV Resonance Raman Studies of the $NaClO_4$ Dependence of Poly-l-lysine Conformation and Hydrogen Exchange Kinetics. *The Journal of Physical Chemistry B*, 116(3):1134–1142, 2012.

[108] Sanford A Asher. Ultraviolet Raman Spectrometry. *Handbook of Vibrational Spectroscopy*, 2002.

[109] Sanford A. Asher. UV Resonance Raman Spectroscopy for Analytical, Physical, and Biophysical Chemistry. Part 1. *Analytical Chemistry*, 65(2):59A–66A, 1993.

[110] H. A. Kramers and W. Heisenberg. Über die Streuung von Strahlung durch Atome. *Zeitschrift für Physik*, 31(1):681–708, 1925.

[111] M. Born, W. Heisenberg, and P. Jordan. Zur Quantenmechanik. II. *Zeitschrift für Physik.*, 35(8):557–615, 1926.

[112] P. A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford, the Clarendon Press, 1930.

[113] David Parker Craig and Thiru Thirunamachandran. *Molecular Quantum Electrodynamics: An Introduction to Radiation-Molecule Interactions*. Courier Corporation, 1984.

[114] Warner L. Peticolas. Multiphoton Spectroscopy. *Annual Review of Physical Chemistry*, 18(1):233–260, 1967.

[115] Gene P. Barnett and A.C. Albrecht. *Raman Spectroscopy: Theory and Practice, Vol. 2*, chapter Appendix: Comments on the Derivation of the Dispersion Equation for Molecules, pages 207–218. Plenum Press, New York, 1970.

[116] G. Placzek. *Handbuch der Radiologie*, volume 6. Akademische Verlagsgesellschaft, Leipzig, 1934.

[117] Andreas C. Albrecht. On the Theory of Raman Intensities. *The Journal of Chemical Physics*, 34(5), 1961.

[118] J. Tang and A. C. Albrecht. Studies in Raman Intensity Theory. *The Journal of Chemical Physics*, 49(3), 1968.

[119] A.B. Myers and R.A. Mathies. *Biological Applications of Raman Spectroscopy, Vol. 2*, chapter Resonance Raman Intensities: a Probe of Excited-State Structure and Dynamics, pages 3–58. Wiley, Canada, 1987.

[120] C. Manneback. Computation of the intensities of vibrational spectra of electronic bands in diatomic molecules. *Physica*, 17(11):1001 – 1010, 1951.

[121] Antonino Natalello, Anna Maria Frana, Annalisa Relini, Alessandra Apicella, Gaetano Invernizzi, Carlo Casari, Alessandra Gliozzi, Silvia Maria Doglia, Paolo Tortora, and Maria Elena Regonesi. A Major Role for Side-Chain Polyglutamine Hydrogen Bonding in Irreversible Ataxin-3 Aggregation. *PLoS One*, 6(4):e18789, 2011.

[122] Xiaoling Wang, Andreas Vitalis, Matthew A. Wyczalkowski, and Rohit V. Pappu. Characterizing the Conformational Ensemble of Monomeric Polyglutamine. *Proteins*, 63(2):297–311, 2006.

[123] Debra S. Caswell and Thomas G. Spiro. Ultraviolet Resonance Raman Spectroscopy of Imidazole, Histidine, and Cu(Imidazole)$_4^{2+}$: Implications for Protein Studies. *J. Am. Chem. Soc*, 108(21):6470–6477, 1986.

[124] Laura M. Markham, Leland C. Mayne, Bruce S. Hudson, and Marek Z. Zgierski. Resonance Raman Studies of Imidazole, Imidazolium, and their Derivatives: the Effect of Deuterium Substitution. *The Journal of Physical Chemistry*, 97(40):10319–10325, 1993.

[125] Kan Xiong, David Punihaole, and Sanford A. Asher. UV Resonance Raman Spectroscopy Monitors Polyglutamine Backbone and Side Chain Hydrogen Bonding and Fibrillization. *Biochemistry*, 51(29):5822–5830, 2012.

[126] X. G. Chen, Reinhard Schweitzer-Stenner, Sanford A. Asher, Noemi G. Mirkin, and Samuel Krimm. Vibrational Assignments of Trans-N-Methylacetamide and Some of Its Deuterated Isotopomers from Band Decomposition of IR, Visible, and Resonance Raman Spectra. *The Journal of Physical Chemistry*, 99(10):3074–3083, 1995.

[127] Alexander V. Mikhonin, Zeeshan Ahmed, Anatoli Ianoul, and Sanford A. Asher. Assignments and Conformational Dependencies of the Amide III Peptide Backbone UV Resonance Raman Bands. *The Journal of Physical Chemistry B*, 108(49):19020–19028, 2004.

[128] Aleksandr V. Mikhonin and Sanford A. Asher. Uncoupled Peptide Bond Vibrations in α-Helical and Polyproline II Conformations of Polyalanine Peptides. *The Journal of Physical Chemistry B*, 109(7):3047–3052, 2005.

[129] T. Shimanouchi. National Bureau of Standards, 1972.

[130] Sanford A. Asher, Richard W. Bormett, X. G. Chen, Donald H. Lemmon, Namjun Cho, Pete Peterson, Marco Arrigoni, Luis Spinelli, and Jeff Cannon. UV Resonance Raman Spectroscopy Using a New cw Laser Source: Convenience and Experimental Simplicity. *Appied Spectroscopy*, 47(5):628–633, 1993.

[131] Sergei Bykov, Igor Lednev, Anatoli Ianoul, Aleksandr Mikhonin, Calum Munro, and Sanford A. Asher. Steady-State and Transient Ultraviolet Resonance Raman Spectrometer for the 193–270 nm Spectral Region. *Applied Spectroscopy*, 59(12):1541–1552, 2005.

[132] D. F. Shriver and J. B. R. Dunn. The Backscattering Geometry for Raman Spectroscopy of Colored Materials. *Applied Spectroscopy*, 28(4):319–323, 1974.

[133] Michael Ludwig and Sanford A. Asher. Self-Absorption in Resonance Raman and Rayleigh Scattering: A Numerical Solution. *Applied Spectroscopy*, 42(8):1458–1466, Nov 1988.

[134] John M. Dudik, Craig R. Johnson, and Sanford A. Asher. Wavelength Dependence of the Preresonance Raman Cross Sections of CH_3CN, SO_4^{2-}, ClO_4^-, and NO_3^-. *The Journal of Physical Chemistry*, 82(4):1732–1740, 1985.

[135] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

[136] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al. Gaussian09 revision D.01. Gaussian Inc. Wallingford CT 2009.

[137] Yan Zhao and Donald G. Truhlar. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-class Functionals and 12 Other Functionals. *Theoretical Chemistry Accounts*, 120(1-3):215–241, 2008.

[138] J.M.L Martin and C. van Alsenoy. GAR2PED: A Program to Obtain a Potential Energy Distribution from a Gaussian Archive Record. University of Antwerp, Antwerp, Belgium, 2009.

[139] K. W. F. Kohlrausch and A. Pongratz. Raman-Effekt und Konstitutions-Probleme, V. Mitteil.: Keto-Enol-Tautomerie in β-Ketosaeure-estern. *Ber. Dtsch. Chem. Ges.*, 67(6):976–989, 1934.

[140] R. Ananthakrishnan. The Raman Spectra of Crystal Powders IV. Some Organic and Inorganic Compounds. *Proceedings of the Indian National Academy of Sciences*, 5(3):200–221, 1937.

[141] J. E. Katon, W. R. Feairheller, and J. V. Pustinger. Infrared Spectra-Structure Correlations of Aliphatic Amides in the 700- to 250-Cm^{-1} Region. *Analytical Chemistry*, 36(11):2126–2130, 1964.

[142] Katsunosuke Machida, Shigeo Kojima, and Toyozo Uno. Infrared Spectra of Finite Chain Molecules–III: N-Fatty Acid Amides. *Spectrochimica Acta, Part A*, 28(2):235 – 256, 1972.

[143] Yoshihiro Kuroda, Yutaka Saito, Katsunosuke Machida, and Toyozo Uno. Vibrational Spectra of Propionamide and Its C- and N-deuterated Compounds. *The Bulletin of the Chemical Society of Japan*, 45:2371–2383, 1972.

[144] G. Nandini and D.N. Sathyanarayana. Ab Initio Studies on Molecular Conformation and Vibrational Spectra of Propionamide. *Journal of Molecular Structure (THEOCHEM)*, 586(13):125 – 135, 2002.

[145] A. Usanmaz and G. Adler. Structure of Propionamide at 123 K. *Acta Crystallographica, Sect. B*, 38(2):660–662, 1982.

[146] David Punihaole, Ryan S. Jakubek, Elizabeth M. Dahlburg, Zhenmin Hong, Nataliya S. Myshakina, Steven Geib, and Sanford A. Asher. UV Resonance Raman Investigation of the Aqueous Solvation Dependence of Primary Amide Vibrations. *The Journal of Physical Chemistry B*, 119(10):3931–3939, 2015.

[147] Leigh B. Clark. Polarization Assignments in the Vacuum UV Spectra of the Primary Amide, Carboxyl, and Peptide Groups. *Journal of the American Chemical Society*, 117(30):7974–7986, 1995.

[148] B. Nolin and R. Norman Jones. The Infrared Absorption Spectra of Diethyl Ketone and its Deuterium Substitution Products. *Journal of the American Chemical Society*, 75(22):5626–5628, 1953.

[149] Lionel John Bellamy. *The Infra-red Spectra of Complex Molecules, Vol. 1* . Chapman and Hall London, 3rd ed. edition, 1975.

[150] J.H. Schachtschneider and R.G. Snyder. Vibrational Analysis of the N-paraffins–II: Normal Co-ordinate Calculations. *Spectrochimica Acta*, 19(1):117 – 168, 1963.

[151] Isao Suzuki. Infrared Spectra and Normal Vibrations of Acetamide and its Deuterated Analogues. *The Bulletin of the Chemical Society of Japan*, 35(8):1279–1286, 1962.

[152] Yoko Sugawara, Akiko Y. Hirakawa, Masamichi Tsuboi, Shigeki Kato, and Keiji Morokuma. Ab Initio SCF MO Study on the Force Field of Amides. *Journal of Molecular Spectroscopy*, 115(1):21 – 33, 1986.

[153] Yoshiyuki Hase. Theoretical Study of the Force Field and Vibrational Assignments of Acetamide and Deuterated Analogues. *Spectrochimica Acta, Part A*, 51(14):2561 – 2573, 1995.

[154] Leland C. Mayne and Bruce Hudson. Resonance Raman Spectroscopy of N-methylacetamide: Overtones and Combinations of the C-N Stretch (Amide II') and Effect of Solvation on the C=O Stretch (Amide I) Intensity. *The Journal of Physical Chemistry*, 95(8):2962–2967, 1991.

[155] Yang Wang, Roberto Purrello, Savas Georgiou, and Thomas G. Spiro. UVRR Spectroscopy of the Peptide Bond. 2. Carbonyl H-Bond Effects on the Ground- and Excited-State Structures of N-methylacetamide. *Journal of the American Chemical Society*, 113(17):6368–6377, 1991.

[156] Laura M. Markham and Bruce S. Hudson. Ab Initio Analysis of the Effects of Aqueous Solvation on the Resonance Raman Intensities of N-Methylacetamide. *The Journal of Physical Chemistry*, 100(7):2731–2737, 1996.

[157] Bruce S. Hudson and Laura M. Markham. Resonance Raman Spectroscopy as a Test of Ab Initio Methods for the Computation of Molecular Potential Energy Surfaces. *Journal of Raman Spectroscopy*, 29(6):489–500, 1998.

[158] Eigil B. Nielsen and John A. Schellman. The Absorption Spectra of Simple Amides and Peptides. *The Journal of Physical Chemistry*, 71(7):2297–2304, 1967.

[159] Eric J. Heller, Robert Sundberg, and David Tannor. Simple Aspects of Raman Scattering. *The Journal of Physical Chemistry*, 86(10):1822–1833, 1982.

[160] Sanford A. Asher, Zhenhuan Chi, and Pusheng Li. Resonance Raman Examination of the Two Lowest Amide $\pi\pi^*$ Excited States. *Journal of Raman Spectroscopy*, 29(10-11):927–931, 1998.

[161] X. G. Chen, Sanford A. Asher, Reinhard Schweitzer-Stenner, Noemi G. Mirkin, and Samuel Krimm. UV Raman Determination of the $\pi\pi^*$ Excited State Geometry of N-Methylacetamide: Vibrational Enhancement Pattern. *Journal of the American Chemical Society*, 117(10):2884–2895, 1995.

[162] Martha A. Nance. Clinical Aspects of CAG Repeat Diseases. *Brain Pathology*, 7(3):881–900, 1997.

[163] Christopher A. Ross and Michelle A. Poirier. Protein Aggregation and Neurodegenerative Disease. *Nature Medicine*, 10:S10–S17, 2004.

[164] Montserrat Arrasate and Steven Finkbeiner. Protein Aggregates in Huntington's disease. *Experimental Neurology*, 238(1):1–11, 2012.

[165] Elyse S. Blum, Andrew R. Schwendeman, and Shai Shaham. PolyQ Disease: Misfiring of a Developmental Cell Death Program? *Trends in Cell Biology*, 23(4):168–174, 2013.

[166] Melissa D. Michelitsch and Jonathan S. Weissman. A Census of Glutamine/Asparagine-rich Regions: Implications for their Conserved Function and the Prediction of Novel Prions. *Proceedings of the National Academy of Sciences*, 97(22):11910–11915, 2000.

[167] Rebecca Nelson, Michael R. Sawaya, Melinda Balbirnie, Anders O. Madsen, Christian Riekel, Robert Grothe, and David Eisenberg. Structure of the Cross-β Spine of Amyloid-like Fibrils. *Nature*, 435:773–778, 2005.

[168] Jed J. W. Wiltzius, Meytal Landau, Rebecca Nelson, Michael R. Sawaya, Marcin I. Apostol, Lukasz Goldschmidt, Angela B. Soriaga, Duilio Cascio, Kanagalaghatta Rajashankar, and David Eisenberg. Molecular Mechanisms for Protein-Encoded Inheritance. *Nature Structural and Molecular Biology*, 16(9):973–8, 2009.

[169] Takashi Miura, Hideo Takeuchi, and Issei Harada. Tryptophan Raman Bands Sensitive to Hydrogen Bonding and Side-chain Conformation. *Journal of Raman Spectroscopy*, 20(10):667–671, 1989.

[170] Gregory P. Harhay and Bruce S. Hudson. Ultraviolet Resonance Raman Study of Proline Isomerization. *The Journal of Physical Chemistry*, 95(9):3511–3513, 1991.

[171] Teruhiko Maruyama and Hideo Takeuchi. Effects of Hydrogen Bonding and Side-chain Conformation on the Raman bands of Tryptophan-2,4,5,6,7-d_5. *Journal of Raman Spectroscopy*, 26(4):319–324, 1995.

[172] Trace Jordan, Ishita Mukerji, Yang Wang, and Thomas G. Spiro. UV Resonance Raman Spectroscopy and Hydrogen Bonding of the Proline Peptide Bond. *Journal of Molecular Structure*, 379(1):51–64, 1996.

[173] Hideo Takeuchi. Raman Spectral Marker of Tryptophan Conformation: Theoretical Basis and Extension to a Wider Range of Torsional Angle. *Journal of Molecular Structure*, 1023(0):143 – 148, 2012.

[174] Nancy E. Triggs and James J. Valentini. An Investigation of Hydrogen Bonding in Amides Using Raman Spectroscopy. *The Journal of Physical Chemistry*, 96(17):6922–6931, 1992.

[175] Lu Ma, Zeeshan Ahmed, Aleksandr V. Mikhonin, and Sanford A. Asher. UV Resonance Raman Measurements of Poly-l-Lysine's Conformational Energy Landscapes: Dependence on Perchlorate Concentration and Temperature. *The Journal of Physical Chemistry B*, 111(26):7675–7680, 2007.

[176] Kan Xiong, Eliana K. Asciutto, Jeffry D. Madura, and Sanford A. Asher. Salt Dependence of an α-Helical Peptide Folding Energy Landscapes. *Biochemistry*, 48(45):10818–10826, 2009.

[177] Kan Xiong and Sanford A. Asher. Circular Dichroism and UV Resonance Raman Study of the Impact of Alcohols on the Gibbs Free Energy Landscape of an α-Helical Peptide. *Biochemistry*, 49(15):3336–3342, 2010.

[178] Zhenmin Hong, Krishnan Damodaran, and Sanford A. Asher. Sodium Dodecyl Sulfate Monomers Induce XAO Peptide Polyproline II to α-Helix Transition. *The Journal of Physical Chemistry B*, 118(36):10565–10575, 2014.

[179] M.V. Shapovalov and R. L. Dunbrack. A Smoothed Backbone-Dependent Rotamer Library for Proteins Derived from Adaptive Kernel Density Estimates and Regressions. *Structure*, 19(6):844 – 858, 2011.

[180] David Punihaole, Zhenmin Hong, Ryan S. Jakubek, Elizabeth M. Dahlburg, Steven Geib, and Sanford A. Asher. Glutamine and Asparagine Side Chain Hyperconjugation-Induced Structurally Sensitive Vibrations. *The Journal of Physical Chemistry B*, 119(41):13039–13051, 2015.

[181] P. Dhamelincourt and F.J. Ramirez. Polarized Micro-Raman and FT-IR Spectra of L-glutamine. *Applied Spectroscopy*, 47:446–451, 1993.

[182] F. J. Ramrez, I. Tun, and E. Silla. Amino Acid Chemistry in Solution: Structural Study and Vibrational Dynamics of Glutamine in Solution. An ab Initio Reaction Field Model. *The Journal of Physical Chemistry B*, 102(32):6290–6298, 1998.

[183] F. Bohlmann. Zur Konfigurationsbestimmung von Chinolizin-Derivaten. *Angewandte Chemie*, 69(20):641–642, 1957.

[184] Jenn-Huei Lii, Kuo-Hsiang Chen, and Norman L. Allinger. Alcohols, Ethers, Carbohydrates, and Related Compounds Part V. The Bohlmann Torsional Effect. *The Journal of Physical Chemistry A*, 108(15):3006–3015, 2004.

[185] Sergei V. Bykov, Nataliya S. Myshakina, and Sanford A. Asher. Dependence of Glycine CH_2 Stretching Frequencies on Conformation, Ionization State, and Hydrogen Bonding. *The Journal of Physical Chemistry B*, 112(18):5803–5812, 2008.

[186] Ramkrishna Adhikary, Jörg Zimmermann, Jian Liu, Ryan P. Forrest, Tesia D. Janicki, Philip E. Dawson, Steven A. Corcelli, and Floyd E. Romesberg. Evidence of an Unusual N–H N Hydrogen Bond in Proteins. *Journal of the American Chemical Society*, 136(39):13474–13477, 2014.

[187] N. H. Rhys, A. K. Soper, and L. Dougan. The Hydrogen-Bonding Ability of the Amino Acid Glutamine Revealed by Neutron Diffraction Experiments. *The Journal of Physical Chemistry B*, 116(45):13308–13319, 2012.

[188] W. Cochran and B. R. Penfold. The Crystal Structure of L-Glutamine. *Acta Crystallographica*, 5(5):644–653, 1952.

[189] Aleksandr V. Mikhonin, Nataliya S. Myshakina, Sergei V. Bykov, and Sanford A Asher. UV Resonance Raman Determination of Polyproline II, Extended 2.5_1-helix, and β-sheet Ψ Angle Energy Landscape in Poly-L-Lysine and Poly-L-Glutamic Acid. *Journal of the American Chemical Society*, 127(21):7712–20, 2005.

[190] Jane S Richardson. *The Anatomy and Taxonomy of Protein Structure*, volume 34. Academic Press, 1981.

[191] Scott A. Hollingsworth, Donald S. Berkholz, and P. Andrew Karplus. On the Occurrence of Linear Groups in Proteins. *Protein Science*, 18(6):1321–1325, 2009.

[192] Laura Mangiarini, Kirupa Sathasivam, Mary Seller, Barbara Cozens, Ale Harper, Colin Hetherington, Martin Lawton, Yvon Trottier, Han Lehrach, Stephen W. Davies, and Gillian P. Bates. Exon 1 of the HD Gene with an Expanded CAG Repeat is Sufficient to Cause a Progressive Neurological Phenotype in Transgenic Mice. *Cell*, 87(3):493–506, 1996.

[193] Eberhard Scherzinger, Rudir Lurz, Mark Turmaine, Laura Mangiarini, Birgit Hollenbach, Renate Hasenbank, Gillian P. Bates, Stephen W. Davies, Hans Lehrach, et al. Huntingtin-Encoded Polyglutamine Expansions Form Amyloid-like Protein Aggregates In Vitro and In Vivo. *Cell*, 90(3):549–558, 1997.

[194] Aneta T. Petkova, Wai-Ming Yau, and Robert Tycko. Experimental Constraints on Quaternary Structure in Alzheimer's β-Amyloid Fibrils. *Biochemistry*, 45(2):498–512, 2006.

[195] Joshua T. Berryman, Sheena E. Radford, and Sarah A. Harris. Thermodynamic Description of Polymorphism in Q- and N-Rich Peptide Aggregates Revealed by Atomistic Simulation. *Biophysical Journal*, 97(1):1 – 11, 2009.

[196] David Punihaole, Riley J. Workman, Zhenmin Hong, Jeffry D. Madura, and Sanford A. Asher. Polyglutamine Fibrils: New Insights into Antiparallel β-Sheet Conformational Preference and Side Chain Structure. *The Journal of Physical Chemistry B.*, 120(12):3012–3026, 2016.

[197] Lauren E. Buchanan, Joshua K. Carr, Aaron M. Fluitt, Andrew J. Hoganson, Sean D. Moran, Juan J. de Pablo, James L. Skinner, and Martin T. Zanni. Structural Motif of Polyglutamine Amyloid Fibrils Discerned with Mixed-Isotope Infrared Spectroscopy. *Proceedings of the National Academy of Sciences*, 111(16):5796–5801, 2014.

[198] Robert Tycko and Reed B. Wickner. Molecular Structures of Amyloid and Prion Fibrils: Consensus versus Controversy. *Accounts of Chemical Research*, 46(7):1487–1496, 2013.

[199] Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc.: Montreal, Canada, 2013.

[200] James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kalé, and Klaus Schulten. Scalable Molecular Dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–802, 2005.

[201] Jing Huang and Alexander D. Mackerell. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *Journal of Computational Chemistry*, 34(25):2135–2145, 2013.

[202] Kresten Lindorff-Larsen, Stefano Piana, Kim Palmo, Paul Maragakis, John L. Klepeis, Ron O. Dror, and David E. Shaw. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins*, 78(8):1950–1958, 2010.

[203] Kresten Lindorff-Larsen, Paul Maragakis, Stefano Piana, Michael P. Eastwood, Ron O. Dror, and David E. Shaw. Systematic Validation of Protein Force Fields Against Experimental Data. *PloS One*, 7(2):e32131, 2012.

[204] W. Humphrey, A. Dalke, and K. Schulten. VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics*, 14(1):33–38, 1996.

[205] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of Simple Potential Functions for Simulating Liquid Water. *The Journal of Chemical Physics*, 79(2):926, 1983.

[206] Tom Darden, Darrin York, and Lee Pedersen. Particle Mesh Ewald: An Nlog(N) Method for Ewald Sums in Large Systems. *The Journal of Chemical Physics*, 98(12):10089–10092, 1993.

[207] Michael R. Shirts and John D. Chodera. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *The Journal of Chemical Physics*, 129(12):1–13, 2008.

[208] Pawel Sikorski, Edward D.T. Atkins, and Louise C. Serpell. Structure and texture of fibrous crystals formed by alzheimer's aβ(11-25) peptide fragment. *Structure*, 11(8):915 – 926, 2003.

[209] William R. W. Welch, Jan Kubelka, and Timothy A. Keiderling. Infrared, Vibrational Circular Dichroism, and Raman Spectral Simulations for β-Sheet Structures with Various Isotopic Labels, Interstrand, and Stacking Arrangements Using Density Functional Theory. *The Journal of Physical Chemistry B*, 117(36):10343–10358, 2013.

[210] Sean D. Moran and Martin T. Zanni. How to Get Insight into Amyloid Structure and Formation from Infrared Spectroscopy. *The Journal of Physical Chemistry Letters*, 5(11):1984–1993, 2014.

[211] T Uni, Katsunosuke Machida, and Yutaka Saito. Infrared Spectra of Partially Deuterated Acetamide. *The Bulletin of the Chemical Society of Japan*, 42:897–904, 1969.

[212] Toyozo Uno, Katsuitosukb Machida, and Yuttaka Saito. Out-of-plane Vibrations of Acetamide and Partially N-deuterated Acetamide. *Spectrochimica Acta, Part A*, 27(6):833 – 844, 1971.

[213] Sven Hovmöller, Tuping Zhou, and Tomas Ohlson. Conformations of Amino Acids in Proteins. *Acta Crystallographica, Section D*, 58(5):768–776, 2002.

[214] Karen E. Marshall and Louise C. Serpell. Insights into the Structure of Amyloid Fibrils. *Open Biology*, 2:185–192, 2009.

[215] Michael R. Sawaya, Shilpa Sambashivan, Rebecca Nelson, Magdalena I. Ivanova, Stuart A. Sievers, Marcin I. Apostol, Michael J. Thompson, Melinda Balbirnie, Jed J.W. Wiltzius, Heather T. McFarlane, and et al. Atomic Structures of Amyloid Cross-β Spines Reveal Varied Steric Zippers. *Nature*, 447(7143):453–457, 2007.

[216] András Perczel, Zoltán Gáspári, and Imre G. Csizmadia. Structure and Stability of β-Pleated Sheets. *Journal of Computational Chemistry*, 26(11):1155–1168, 2005.

[217] Mathur R. Narasimhamurthy, Kailasam Venkatesan, and Fritz Winkler. Crystal and molecular structure of N-acetyl–glutamine. *Journal of the Chemical Society, Perkin Transactions 2*, pages 768–771, 1976.

[218] K. Panneerselvam and M. Soriano-García. L-Glycyl-L-Glutamine Monohydrate. *Acta Crystallographica, Section C*, 51(12):2718–2720, 1995.

[219] Carl Henrik Görbitz and Lars Male Hartviksen. The monohydrates of the four polar dipeptides L-seryl-L-asparagine, L-seryl-L-tyrosine, L-tryptophanyl-L-serine and L-tyrosyl-L-tryptophan. *Acta Crystallographica, Section C*, 64(3):o171–o176, 2008.

[220] George M. Sheldrick. A short history of *SHELX*. *Acta Crystallographica, Section A*, 64(1):112–122, 2008.

[221] Nighat Kausar, Bruce D. Alexander, Trevor J. Dines, Robert Withnall, and Babur Z. Chowdhry. Vibrational spectroscopy and dft calculations of amino acid derivatives: N-acetyl-l-asp and n-acetyl-l-glu in the solid state. *Journal of Raman Spectroscopy*, 40(6):670–678, 2009.

[222] Sujit Banerjee and Dongyun Li. Interpreting Multicomponent Infrared Spectra by Derivative Minimization. *Applied Spectroscopy*, 45(6):1047–1049, 1991.

[223] Michael A. Friese and Sujit Banerjee. Lignin Determination by FT-IR. *Applied Spectroscopy*, 46(2):246–248, 1992.

[224] Zhenhuan Chi, X. G. Chen, Janet S. W. Holtz, , and Sanford A. Asher. UV Resonance Raman-Selective Amide Vibrational Enhancement: Quantitative Methodology for Determining Protein Secondary Structure. *Biochemistry*, 37(9):2854–2864, 1998.

[225] Noemi G. Mirkin and Samuel Krimm. Amide III Mode ϕ, ψ Dependence in Peptides: A Vibrational Frequency Map. *The Journal of Physical Chemistry A*, 106(14):3391–3394, 2002.

[226] Charles H. Bennett. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *Journal of Computational Physics*, 22(2):245 – 268, 1976.